PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

A 22

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 : C12N 15/00	A2	(11) International Publication Number: WO 00/65039
		(43) International Publication Date: 2 November 2000 (02.11.00)

(21) International Application Number: PCT/US00/11460

(22) International Filing Date: 28 April 2000 (28.04.00)

(30) Priority Data:

60/131,433	28 April 1999 (28.04.99)	US
60/144,669	19 July 1999 (19.07.99)	US

(71) Applicant *(for all designated States except US)*: NOVARTIS AG [CH/CH]; Schwarzwaldallee 215, CH–4058 Basel (CH).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: MALECK, Klaus [DE/DE]; Augustenstrasse 74, D–80333 München (DE). LAWTON, Kay, Ann [US/US]; 8016 Crooked Chute Court, Raleigh, NC 27612 (US). DIETRICH, Robert, Arthur [US/US]; 25 Copper Hill Court, Durham, NC 27713 (US).

(74) Agent: MEIGS, J., Timothy; Novartis AG, Patent & Trademark Dept., Agribusiness, Site Rosental, CH–4002 Basel (CH).

(81) Designated States: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published

Without international search report and to be republished upon receipt of that report.

(54) Title: IDENTIFICATION OF DNA CONTROL ELEMENTS RESPONSIVE TO SPECIFIC STIMULI

(57) Abstract

In order to identify genes whose proteins are involved in the regulation of SAR downstream of cell death, a screen was carried out to isolate mutants that constitutively express SAR. The *PR–1* promoter, the most reliable marker for the onset of SAR, was cloned in front of the firefly *(Photinus pyralis)* luciferase reporter gene and transformed into *Arabidopsis*. A transgenic line with a luciferase expression pattern that paralleled expression of the endogenous *PR–1* gene was identified and subjected to EMS mutagenesis to isolate mutants that constitutively expressed the *PR–1*/luciferase gene. Biochemical, cytological, pathological, and genetic methods were employed to further characterize the mutants and to prove the isolation of SAR mutants downstream of cell death. This characterization enables one to distinguish between classes of *cim* mutants, and to describe physiological changes that occur during the maintenance phase of SAR. Furthermore, utilizing microarray chips, the entire plant genome was simultaneously surveyed for genes that change in expression in response to biotic and abiotic factors. By comparing gene expression changes across various treatments, groups of co–regulated genes (regulons) were identified and the genomic sequences of genes within a regulon were examined to identify common sequence motifs that are likely to act as regulatory elements. This approach used experimental design based upon the biology of the study system in combination with bioinformatics to analyze the results.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

IDENTIFICATION OF DNA CONTROL ELEMENTS RESPONSIVE TO SPECIFIC STIMULI

FIELD OF THE INVENTION

The invention generally relates to the use of gene expression profiling to identify groups of genes that show a similar pattern of expression in response to biotic and abiotic stimuli. The invention more particularly relates to the use of the expression profile data thus generated to uncover gene groups that are co-regulated and to identify common DNA sequences that function to regulate gene expression in response to specific stimuli.

BACKGROUND OF THE INVENTION

The SAR Signal Transduction Cascade:

Plants are constantly challenged by a wide variety of pathogenic organisms including viruses, bacteria, fungi, and nematodes. Crop plants are particularly vulnerable because they are usually grown as genetically-uniform monocultures; when disease strikes, losses can be severe. However, most plants have their own innate mechanisms of defense against pathogenic organisms. Natural variation for resistance to plant pathogens has been identified by plant breeders and pathologists and bred into many crop plants. These natural disease resistance genes often provide high levels of resistance to or immunity against pathogens.

Systemic acquired resistance (SAR) is one component of the complex system plants use to defend themselves from pathogens (Hunt and Ryals, 1996; Ryals *et al.*, 1996). *See also,* U.S. Patent No. 5,614,395. SAR is a particularly important aspect of plant-pathogen responses because it is a pathogen-inducible, systemic resistance against a broad spectrum of infectious agents, including viruses, bacteria, and fungi. When the SAR signal transduction pathway is blocked, plants become more susceptible to pathogens that normally cause disease, and they also become susceptible to some infectious agents that would not normally cause disease (Gaffney *et al.*, 1993; Delaney *et al.*, 1994; Delaney *et al.*, 1995; Delaney, 1997; Bi *et al.*, 1995; Mauch-Mani and Slusarenko, 1996). These observations indicate that the SAR signal transduction pathway is critical for maintaining plant health.

Induction Of SAR:

SAR is triggered by certain necrosis, as found as part of disease symptoms or the hypersensitive response (HR). HR is a genetically programmed cell death process that is induced in many incompatible plant-pathogen interactions. It is also simulated in certain lesion mimic mutants in absence of pathogens (Dangl *et al.*, 1996). These mutants, called *lsd* (lesions simulating disease resistance), *acd* (accelerated cell death) or *cpr* (constitutive PR genes) are in many cases also constitutively activated for SAR (Bowling *et al.*, 1994; Dietrich *et al.*, 1994; Greenberg and Ausubel, 1993). Lesion mimic phenotypes can be caused by mutations in *R* genes, as seen in the *Rp1* mutant in maize (Hu *et al.*, 1996) but they can also be caused by metabolic perturbations (Dangl *et al.*, 1996), and loss-of-function mutations in putative transcription factors (e.g. *lsd1*; Dietrich *et al.*, 1997). Mutants deficient for R gene-specific signal transduction provide another link between the induction of HR and SAR. The mutant *ndr1* (non-race specific disease resistance), which was isolated in a screen for susceptibility to avirulent Pseudomonas syringae strains, is deficient for the induction of a local defense reaction induced by *R* genes of the LZ-NBS-LRR class (Aarts *et al.*, 1998; Century *et al.*, 1995). Subsequently, SAR cannot be induced in these interactions in the *ndr1* mutant. Similarly, in the *eds1* (enhanced disease susceptibility) mutant, the signaling cascade induced by *R* genes containing an N-terminal TIR domain is interrupted (Aarts *et al.*, 1998; Parker *et al.*, 1996). As a consequence, in interactions involving TIR-NBS-LRR R genes, no SAR is induced in *eds1*. In these two mutants, the block in the induction of the SAR response may actually reside in the suppression of the formation of an HR rather than in a block of the SAR signaling cascade itself (Table 1 summarizes information on *Arabidopsis* mutants in disease resistance).

Table 1: Disease Resistance Mutants in *Arabidopsis thaliana* (adapted from Maleck and Lawton, 1998)

Mutant	Screen and possible function	Reference
***Avr-R* gene interaction**		
eds1 (enhanced disease susceptibility)	Susceptibility to avirulent *Peronospora parasitica* isolates, member of the converging TIR-NBS-LRR R gene signaling pathway	(Parker *et al.*, 1996)
ndr1 (non-race-specific disease resistance)	Susceptibility to avirulent *Pseudomonas syringae* strains, convergence of LZ-NBS-LRR R gene signaling	(Century *et al.*, 1995)
lsd1, 6 (lesions simulating disease resistance)	Identification of spontaneous lesion formation. Wild-type alleles are involved in limiting initiation or spreading of cell death	(Dietrich *et al.*, 1994)
acd2 (accelerated cell death)	Same as *lsds*	(Greenberg and Ausubel, 1993)
cell death		
cims/cpr (constitutive immunity/ constitutive *PR* gene expression)	Marker gene overexpression (*PR-1* or *PR-2*); Role in SA biosynthesis or SAR upregulation	(Bowling *et al.*, 1994)
dnd1 (defense, no death)	Absence of HR when inoculated with avirulent *Pseudomonas syringae*, constitutive immunity	(Yu *et al.*, 1998)
pad (phytoalexin deficient)	No phytoalexin accumulation after infection by the moderate virulent pathogen *Pseudomonas syringae* pv. *maculicola* ES4326. Genes may be involved in phytoalexin biosynthesis or in general pathogen recognition and signaling	(Glazebrook and Ausubel, 1994)
SA accumulation		
nim1/npr1/sai1 (no immunity/no PR genes/SA insensitive)	Susceptibility to virulent *Peronospora parasitica* isolates after chemical immunization, hypersensitive to *Pseudomonas syringae*, counter selection using a SA-inducible promoter; Nim1 seems to be a central component of SAR	(Cao *et al.*, 1994; Delaney *et al.*, 1995; Shah *et al.*, 1997)
PR gene expression		
non-SAR mutants		
edr1 (enhanced disease resistance)	Resistance to virulent *Pseudomonas syringae* pathovars; also resistant to *Erisyphe cichoracearum*	(Frye and Innes, 1998)

SAR can be distinguished from other inducible disease resistance responses by a number of associated biochemical and physiological changes, which eventually confer an increased resistance to a secondary pathogen attack.

SAR Marker Proteins:

The identification of SAR specific marker proteins was crucial to distinguish SAR from other inducible plant resistance responses and to allow its dissection. SAR markers are proteins whose expression is tightly correlated with the onset of SAR in uninfected tissue (Metraux *et al.*, 1989; Uknes *et al.*, 1992; Ward *et al.*, 1991). All of the SAR markers fall in the class of PR proteins, which were originally identified as low-molecular weight, acidic proteins accumulating after TMV infection of tobacco leaves (Van Loon, 1985) or other pathological and stress-related situations (Kombrink and Somssich, 1997). In tobacco, the set of SAR markers is encoded by at least nine gene families (Alexander *et al.*, 1992; Ward *et al.*, 1991). In *Arabidopsis*, three SAR markers have been identified, called PR-1, PR-2 and PR-5 (Uknes *et al.*, 1992), with PR-1 being the most specific marker protein. The *Arabidopsis* PR-1 promoter has been examined in detail for responsiveness to salicylic acid (SA) and 2,6-dichloro-isonicotinic acid (INA). Only two or three active promoter domains have been identified, none of them with homology to ethylene response binding protein sites, which may reflect on a molecular level the observed overall specificity of PR-1 induction (Lebel *et al.*, 1998).

SAR proteins accumulate to high levels during the maintenance phase of systemic acquired resistance, suggesting that these proteins contribute to resistance. In support of this idea, in vitro antimicrobial activity has been described for tobacco PR-1a and other PR proteins. In vivo studies involving overexpression of PR-1a in tobacco have demonstrated a significant increase in resistance to infection by the two oomycete pathogens, *P. tabacina* and *P. parasitica* var. *nicotianae* (Alexander *et al.*, 1993). This result, and similar findings with other SAR proteins, suggests that these proteins are causally associated with disease resistance. While the function of the PR-2 and PR-5 proteins can be derived from their sequences as β-1,3-glucanases and thaumatin/osmotin-like proteins, respectively, the function of the PR-1 protein remains elusive.

The Role Of Salicylic Acid In SAR:

Early studies revealed that exogenous application of salicylic acid (SA) or acetylsalicylic acid could induce disease resistance and the expression of PR genes (Van Loon and Antoniw, 1982; White, 1979). In successive experiments, it was shown that SA accumulation is not only concomitant to but also required for SAR induction. SA is synthesized in plant cells from phenylalanine, which is converted to t-cinnamic acid (t-CA) by PAL, a key enzyme in the phenylpropanoid pathway. t-CA is converted to SA via benzoic acid, presumably in the chloroplast (Yalpani *et al.*, 1993). Free SA is rapidly conjugated to the β-*O*-D glycoside (Enyedi *et al.*, 1992) by an SA-inducible UDP-glucose:SA 3-*O*-glycosyltransferase (Enyedi and Raskin, 1993). Only free SA is active, but its action can be mimicked by the functional analogs BTH (Benzothiadiazol, CGA 245704) and INA (2,6-dichloro-isonicotinic acid, CGA 41396) (Friedrich *et al.*, 1996; Görlach *et al.*, 1996; Lawton *et al.*, 1996; Vernooij *et al.*, 1995). Using the *Pseudomonas putida NahG* gene, which encodes for a salicylate hydroxylase, plants unable to accumulate SA were generated (Gaffney *et al.*, 1993). These plants were also unable to raise an SAR response, and were hypersusceptible to many virulent and avirulent pathogens (Delaney *et al.*, 1994). In grafting experiments with *NahG*-expressing tobacco plants (NahG), it has been shown that SAR induction does not occur in a NahG scion grafted on a wild-type rootstock infected with TMV. But the signal can be produced and transmitted from a *NahG*-expressing rootstock after infection with TMV to a wild-type scion as determined by SAR activation (Vernooij *et al.*, 1994). These results confirmed previous kinetic studies in cucumber that indicated SA is not the signal responsible for systemic induction of SAR (Rasmussen *et al.*, 1991). Furthermore, Willits *et al.* (Willits and Ryals, 1998) showed that, in tobacco, resistance and SAR gene expression are linearly correlated to the SA content. This finding negates the argument that residual SA in NahG plants might be sufficient to activate SAR and leads to a model in which SA is required for defense gene expression in systemic tissue downstream of the systemic signal.

In crosses of the NahG plant to lesion mimic mutants, two different classes of lsd mutants were identified. In the mutants *lsd2* and *lsd4*, lesion formation in these two mutants in a NahG background is unchanged, but the activation of SAR requires previous accumulation of SA (Hunt *et al.*, 1997). In the second class of lesion mimics, comprised of mutants *lsd1*,

lsd6 and *lsd7*, cell death (and SAR gene expression) is suppressed in a NahG background but can be reinduced by exogenous application of INA or BTH, thus positioning cell death and resistance both upstream and downstream of SA (Weymann *et al.*, 1995; R. Dietrich and J. Dangl, pers. comm.). HR and SAR induction appear to be tightly intertwined.
Other pathogen-derived signals, such as cell-wall fragments, elicitors and wounding induce other systemic defense responses in the plant.

Other Inducible Defense Responses And Their Relationship To SAR:

The wound-inducible, jasmonic acid (JA) and ethylene mediated defense response has primarily been studied in the context of induced resistance to insect predation in tomato and tobacco, and more recently, also in *Arabidopsis* (McConn *et al.*, 1997). It is triggered by wounding and insect feeding and results in the induction of proteinase inhibitor (PI) genes (Creelman and Mullet, 1997; Ryan, 1990). The proteinase inhibitors interfere with digestion in the insect gut and discourage further feeding. Evidence is accumulating that the wound response signaling cascade may also play a role in defense against certain fungal pathogens (Staswick *et al.*, 1998; Vijayan *et al.*, 1998). It has been known for some time that the response can also be triggered by fungal elicitors (Boller, 1991). In addition, the signaling molecules jasmonate and ethylene regulate not only wound-response genes, but are also involved in regulating expression of some (non-SAR) PR genes (Reymond and Farmer, 1998). They include genes for PR-3 (chitinase), PR-4 (thaumatin-like protein; Lawton *et al.*, 1994), defensins (PDF1.2 = PR-13; Penninckx *et al.*, 1996; Penninckx *et al.*, 1998) and thionins (= PR-12; Vignutelli *et al.*, 1998) in *Arabidopsis*. Defensins and thionins are small, cystein-rich peptides (~5 kDa) with potent in vitro activity inhibiting fungal growth (Bohlmann, 1994; Broekaert *et al.*, 1995). Structurally related peptides can be found not only in other plant species, but also in insects where they are also likely to participate in antimicrobial host defenses (Hancock *et al.*, 1995). Overexpression of defensins or thionins in plants leads to enhanced resistance to certain pathogens that are not affected by SAR (Epple *et al.*, 1997; Terras *et al.*, 1995).

A genetic block in the wound-response signaling cascade renders plants more susceptible to necrotrophic fungal pathogens. For example, the jasmonate insensitive mutant jar1 and the jasmonate-deficient triple *fad3-2fad7-2fad8* mutant are both more susceptible

than wild-type plants to *Botrytis cinerea* or *Pythium irregulare* but not to "classical" SAR pathogens such as *Peronospora* (Staswick *et al.*, 1998; Thomma *et al.*, 1998; Vijayan *et al.*, 1998). Similarly, mutants in the ethylene perception (*ein* and *etr* in *Arabidopsis*, ectopic *etr* in tobacco) show increased susceptibility (or increased tolerance) to avirulent and virulent pathogens (Bent *et al.*, 1992; Knoester *et al.*, 1998; Lund *et al.*, 1998).
Furthermore, expression of an antisense construct of a lipoxygenase gene in tobacco results in reduced jasmonate synthesis and increased susceptibility to *Rhizoctonia solani* and *Phytophtora parasitica* (Rance *et al.*, 1998). Thus, some components of the wound response may be involved in resistance to certain pathogens as well as in resistance to insects. The pathogen spectrum might however vary from plant to plant and may overlap with the pathogen spectrum of SAR.

These two induced "immune responses" seem to make the major contribution to the overall systemic response. However, other induced resistance pathways have been described. Induced Systemic Resistance (ISR) can be triggered in certain hosts, including *Arabidopsis*, by the biocontrol bacteria *Pseudomonas fluorescens* WCS 417, by *Serratia macescens* or by cell wall preparations of these microorganisms (i.e. lipopolysaccharides) (Pieterse *et al.*, 1996; Van Wees *et al.*, 1997). ISR confers quantitative resistance to fungal (*Fusarium oxysporium*) and bacterial pathogens (*Pseudomonas syringae*) and appears to be independent of SA, but dependent on ethylene and jasmonic acid (Pieterse *et al.*, 1998). Interestingly, none of the typical marker genes for JA/ethylene induced resistance is expressed during ISR.

Similarly, in the *edr1* mutant (enhanced disease resistance), none of the known marker genes is induced while the mutant exhibits a good resistance to *Erysiphe pathovars*, but not to *Peronospora* (Frye and Innes, 1998).

Finally, in tobacco, the bacteria *Erwinia carotovora* induces a systemic resistance that is antagonistic to the induction of SAR. This resistance can be followed by the transcription level of basic PR genes (Vidal *et al.*, 1997). The pathogen specificities and interactions between the different pathways are just emerging (for review, see Maleck and Dietrich, 1999).

Genetic Dissection Of SAR:

To further dissect the signaling downstream of cell death, a screen for mutants unable to activate SAR after INA treatment was performed. Six alleles of a mutant hypersensitive to

Peronospora were isolated (Delaney *et al.*, 1995). In several other screens, including a marker gene based screen using a *PR-2*/GUS reporter gene construct, the same gene was isolated, thus underlining the central importance of the *NIM1/NPR1* (*no immunity*, *no PR* gene expression) gene for SAR activation downstream of SA (Cao *et al.*, 1994). The *nim1/npr1* gene was cloned by map-based cloning. The sequence has homologies to some ankyrin-containing transcription factor regulators, such as IκBα (Cao *et al.*, 1997; Ryals *et al.*, 1997). Overexpression of the *NIM1* gene results in plants that are poised to respond faster and stronger than wild-type plants after (subclinical) concentrations of chemical inducers (Cao *et al.*, 1998; U.S. Patent No. 6,031,153). Interestingly, *NIM1* is also required for the SA-independent ISR (Pieterse *et al.*, 1998). Furthermore, *nim1/npr1* does not entirely block the resistance observed in some of the *cpr* mutants (Bowling *et al.*, 1997; Clarke *et al.*, 1998). For instance, in the mutant *cpr5*, resistance to *Peronospora parasitica* is *nim1*-independent, whereas resistance to *Pseudomonas syringae* is abolished in the *nim1/npr1* background (Bowling *et al.*, 1997). Clearly, SAR is part of a complex signaling network and the expression of SAR in the plant is strictly regulated.

Arabidopsis thaliana Is A Model System To Study SAR:

Because the study of plant-pathogen interactions involves two living organisms, the choice of a model organism is crucial to reduce the complexity of the investigation. Tobacco was among the first plants in which a systemic activation of defense is observed. In 1961, Ross used the tobacco mosaic virus as both the inducing and the challenging pathogen to define SAR and Local Acquired Resistance (LAR) (Ross, 1961a; Ross, 1961b). Tobacco, because of its amenability to biochemical analysis and its facile transformability, remains an important system for the study of SAR. A wealth of pathogens, as well as elicitors that induce SAR in tobacco have been identified (Keller *et al.*, 1996).

Tobacco is however not easily amenable to genetic studies for gaining molecular insights into the components of the SAR signaling cascade. For this approach, research has focused on the "useful weed" *Arabidopsis thaliana* (L.) Heynh. (Meyerowitz, 1989). *Arabidopsis* is attractive as a research tool because of its diminutive stature, short generation time (6 to 8 weeks), high seed yield and its small, well-characterized genome, which makes it of great use in the dissection of other signal transduction pathways, such as the ethylene or the

ABA signal transduction. In addition, members of several important plant pathogens are virulent on *Arabidopsis* (Meyerowitz and Somerville, 1994). Phylogenetically *Arabidopsis* belongs to the Brassicaceae family, which encompasses many crop plants, such as cabbage and mustard (Price *et al.*, 1994).

Molecular Tools In *Arabidopsis*

Arabidopsis thaliana has a small genome (110 Mb) with a high gene density (about 1 gene per 4 kb, or an estimated 21,000 – 25,000 genes in total). Less than 10% of the genome, including centromeric and telomeric repeats, contains dispersed repetitive elements (Goodman *et al.*, 1995; Pruitt and Meyerowitz, 1986). These characteristics make *Arabidopsis* an ideal plant for genetic and physical mapping projects. A large number of genetic markers (Table 2) and genetically diverse land races exist. A significant fraction of the genome has been assembled into physical contigs in high capacity cloning vectors, such as YACs, BACs and P1 clones (Schmidt, 1998). A multinational sequencing effort is expected to finish the deciphering of all nucleotides of the plant's genome by the end of the year 2001 (Ecker, 1998). Already an extensive amount of sequence information is available in the form of genomic sequence, cDNA sequence or as expressed sequence tags (*est*). A consortium of several French universities, the University of Minnesota and Michigan State University (MSU) has accumulated a public collection currently containing over 37,000 clones (Dec. 31 1998), which are derived from estimated 14,000 distinct genes (Delseny *et al.*, 1997; Newman *et al.*, 1994). These resources, along with the ease of mutagenesis, transformation, and tagged mutant collections, have made *Arabidopsis* the model plant for functional genomics (Stiekema and Pereira, 1998).

Table 2: Genetic Marker Types Currently Used In *Arabidopsis* Genetic Mapping Projects (partly derived from Rafalski *et al.*, 1996, p.72).

	RFLP	**RAPD**	**SSLP**	**CAPS**	**AFLP**	**Heteroduplex**
Principle	Genomic DNA digested, blotted and hybridized	PCR with random primers	PCR of simple sequence repeat regions	Digest of PCR products	PCR of a subset of restriction fragments from extended adapter primers	Altered gel migration of heteroduplex PCR fragment
Nature of polymorphism	Single base pair, insertion, deletion	Single base pair, insertion, deletion	Repeat length changes	Single base pair, insertion, deletion	Single base pair, insertion, deletion	Single base pair, insertion, deletion
Abundance in the genome	High	Very high	Medium	High	High	Very high
DNA sequence information	No	No	Yes	Yes	No	Yes
Dominance	codominant	dominant	codominant	codominant	mixed	mixed
DNA required	2 - 10 mg	10 - 25ng	50 - 100 ng	50 - 100 ng	1 - 2 mg	10 - 100 ng
Available markers	>550	N/A[1]	50	120	395	19
References	(Liu *et al.*, 1996) and references therein http://nasc.life.nott.ac.uk/new_ri_map.html	(Reiter *et al.*,1992)	(Bell and Ecker, 1994); http://cbil.humgen.upenn.edu/~atgc/SSLP_info/SSLP.html	(Konieczny and Ausubel, 1993); http://genome-www.stanford.edu/*Arabidopsis*/aboutcaps.html	(Alonso-Blanco *et al.*, 1998)	(Hauser *et al.*, 1998)

1. N/A: not applicable (per primer, several fragments are obtained, with numbers varying with the PCR conditions)

The term "functional genomics" describes the combined efforts to elucidate the functions of the increasing numbers of unknown genes identified by mass-sequencing. Currently, the functions of roughly 50% of all putative genes are unknown (Bevan *et al.*, 1998). Homology searches, expression profiling, knock-out mutant analysis, overexpression studies, and protein-protein interaction analysis might all yield clues to the biochemical, cellular, adaptive or developmental role of a given protein (Bouchez and Hofte, 1998). To match the high-throughput sequencing effort, highly paralleled technologies have been conceived, mostly to gather gene expression patterns (Table 3). Considering the importance of transcriptional changes during plant defense (for example induction of genes such as 4-coumarate:CoA ligase (*4CL*), phenylalanine ammonia lyase (*PAL*), and pathogenesis related (*PR*) genes; Hahlbrock and Scheel, 1989), gene expression profiles potentially play an important role in unraveling the complexity of regulatory pathways in plant-pathogen interactions. Addition references of note include Cho, *et al.* (1998a); Cho, et. al. (1998b); DeRisi, *et al.* (1997); Eisen, *et al.* (1998); Fambrough, *et al.* (1999) ; Gerhold, *et al.* (1999); Iyer, *et al.* (1999); Kehoe, *et al.*(1999); Lemieux, *et al.* (1998); Roth, *et al.* (1998); Schena, *et al.* (1998); Tavozoie, *et al.* (1999); Tomayo, *et al.* (1999); and Winzeler, *et al.* (1999).

Table 3: Survey Of Technologies Of Mass-Gene Expression Profiling

Technology	Principle	Comments	References
SAGE (serial analysis of gene expression)	Short cDNA sequence tags (9 - 11 bp) are concatenated and sequenced	Rarer transcripts may be missed, the genome sequence (or exhaustive *est* sequences) should be available to identify the tags. Not convenient for mass comparisons	(Velculescu *et al.*, 1997)
DNA oligonucleotide array (chip)	For every gene, several short, complementary oligonucleotides (25mers) are photolithographically synthesized on a glass support. Hybridization with fluorescently labeled cDNA populations	To chose the best oligonucleotides for every gene, the entire genome sequence must be known. The nec plus ultra in miniaturization (4,000,000 probes per 1.28 cm^2)	(Wodicka *et al.*, 1997)
DNA microarray	cDNA or *est* fragments are bound to a support (glass or membrane). Either competitive or simple hybridization of labeled cDNA populations is measured	Many different formats are available. Only known genes can be monitored, rare transcripts are likely to be missed.	(Ruan *et al.*, 1998; Schena *et al.*, 1995; Desprez *et al.*, 1998)
cDNA AFLP	gel-based comparison of two populations of cDNAs, labeled by the AFLP technology.	Labor-intensive identification of bands by excision and cloning, not automatable	(Bachem *et al.*, 1996)
Differential display, PCR select	PCR-based normalization and enrichment of differentially abundant RNA molecules, yielding a cDNA library	Leads primarily to the identification of differentially expressed genes, successive expression analysis still to be done. Not high-throughput	(Diatchenko *et al.*, 1996)

Inducible Gene Expression:

A principal advantage to be realized through genetic engineering of plants is the controlled expression of chimeric genes so that they are expressed only at the appropriate time, to the appropriate extent, and in some situations in particular parts of the plant. For example, the energy expended by a plant to continuously produce high levels of a foreign protein could prove detrimental to the plant, whereas if the gene were expressed only when desired, the drain on energy and therefore yield could be reduced. Additionally, the phenotype expressed by the chimeric gene could result in adverse effects to the plant if expressed at inappropriate times during development. For tissue in culture or in a bioreactor, the untimely production of a desired secondary product could lead to a decrease in the growth rate of the culture, resulting in a decrease in the yield of the product. Thus, specific regulation of plant gene expression by exogenous application of chemicals to increase or decrease expression of transgenes of interest could be of particularly great commercial value to both seed and crop protection businesses as well as to end users (e.g. food processors) of agricultural commodities. Promoters activated by specific stimuli can be used for regulated expression of value added traits, input traits and output traits as well as for production of certain proteins (e.g. antibodies, etc).

In view of such considerations, it is apparent that control of the time, extent, and/or site of expression of chimeric genes in plants or plant tissues would be highly desirable. An ideal situation would be the at-will regulation of expression of an engineered trait via a regulating intermediate that could be easily applied to field crops, ornamental shrubs, bioreactors, etc.

Several plant genes are known to be induced by various internal and external factors including plant hormones, heat shock, chemicals, pathogens, lack of oxygen, and light. For example, as discussed above, exogenous application of SA induces SAR and expression of PR genes (Ward, *et al.* 1991; Uknes, *et al.*, 1992) as well as of synthetic compounds such as 2,6-dichloroisonicotinic acid (INA) (Vernooij, *et al.*, 1995) and benzo(1,2,3)thiadiazole-7-carbothioic acid S-methyl ester (BTH) (Friedrich, *et al.*, 1996; Lawton, *et al.*, 1996). Therefore, induction of PR protein genes by chemicals or pathogens provides a foundation to address the problem of controlling gene expression in plants and plant tissue.

Studies of the effects of chemical regulators on the promoters of several PR genes have been described in the literature, shedding additional light of the SAR signal transduction

pathway. Deletion analysis of the tobacco PR-1a promoter revealed that a 600-bp long promoter lost its functional inducibility by pathogen infection as well by exogenous chemical application while 661-bp retained inducibility although to a lesser extent than a 903-bp long fragment (Uknes, *et al.*, 1993). Analysis of the tobacco PR-2d promoter revealed that some of its inducility is lost in a 607-bp long fragment but that 1047-bp are required for maximal induction, while a 321-bp long promoter had lost almost any inducibility (Hennig, *et al.* 1993). Also, a Myb-like transcrition factor (myb1) was isolated and its expression shown to be inducible by SA and tobacco mosaic virus (Yang, *et al.* 1996). Furthermore, it was shown to bind *in vitro* to a fragment of the tobacco PR-1a promoter (positions -679 to -487 from the transcription start site) containing a Myb-like recognition site (positions -520 to -514). Moreover, a sequence in the tobacco PR-2d promoter (-348 to -324) was shown to bind *in vitro* to another protein. Mutation of this sequence reduced inducibility by SA by approximately 3-fold but not completely compared to wild-type sequence when included in a fragment spanning positions -364 to -288 in the PR-2d promoter and fused to the -90 35S CaMV promoter (Shah, *et al.* 1996).

United States Patent No. 5,614,395 describes the *Arabidopsis* PR-1 protein gene and its chemically inducible promoter. As described in this patent, the full-length *Arabidopsis* PR-1 promoter fragment was fused to the firefly luciferase (LUC) gene and ultimately cloned into plasmid pAtPR1-S, which is in turn transformed into *Arabidopsis* plants for chemical induction analysis. The transgenic *Arabidopsis* lines carrying the PR-1 promoter/LUC gene fusion are then treated by spraying with isonicotinic acid (INA). When analyzed, the transgenic lines showed significantly higher induction of luciferase activity compared to water-treated controls. Thus, INA was shown to induce expression in transformed plants of a chimeric gene comprising the full-length *Arabidopsis* PR-1 promoter fragment. Subsequently, WO 98/03536 described deletion mutants of the *Arabidopsis* PR-1 promoter that are shorter than the full-length *Arabidopsis* PR-1 promoter sequence, yet still yield similar induction of gene expression upon the application of a chemical regulator.

Despite the above successes, previous methods for identifying regulatory elements in genes have largely relied upon extensive molecular manipulations commonly referred to as 'promoter bashing' that involve deletion and/or replacement of DNA sequences in the 5' regulatory region of the gene, ligation of the modified promoter sequence to a reporter gene

(e.g. CAT, GUS, luciferase, GFP), transformation of plants and evaluation of effects of the modification on gene expression by monitoring the activity of the reporter gene in the transgenic plants. Such work can be time consuming, and typically only a few genes can be evaluated at a time.

In view of the above, there is a long-felt but unfulfilled need for new methods of discovering chemically inducible regulatory elements, particularly chemically inducible regulatory elements in plants.

SUMMARY OF THE INVENTION

The present invention addresses the aforementioned needs by providing a new method whereby the entire genome (transcriptome) is simultaneously surveyed for genes that change in expression in response to biotic and abiotic factors. By comparing gene expression changes across various treatments, groups of co-regulated genes (regulons) are identified and the genomic sequences of genes within a regulon are examined to identify common sequence motifs that are likely to act as regulatory elements. These regulatory elements are then used to make promoters that drive controlled gene expression. This approach uses experimental design based upon the biology of the study system in combination with bioinformatics to analyze the results.

In a preferred embodiment, gene expression profiling using DNA microarrays is used to study the transcriptome of a plant to identify groups of genes that show a similar pattern of expression in response to biotic and abiotic stimuli, especially biotic and abiotic inducers of SAR. Expression profile data can uncover gene groups that are co-regulated (regulons), and can be used to identify common DNA sequences that function to specifically regulate gene expression in response to exogenous factors but not endogenous signals. The resulting regulatory sequence elements can be cloned and used to precisely regulate genes of interest in transgenic plants.

The present invention is useful for identifying genes that are responsive to BTH and/or pathogens, using PCR select and microarray gene chip technology. Experiments are conducted to compare expression profiles in response to biotic and abiotic inducers of SAR and to assess the requirement for salicylic acid and the *NIM1* gene for mRNA accumulation.

By analysis of data generated with cDNA microarrays, sets of genes that are responsive specifically to exogenous application of BTH, a chemical that can activate the SAR response, can be identified. That there exists a set of BTH-inducible genes that are not responsive to SA or pathogens is especially surprising given the expectation from previous studies of the SAR signal transduction pathway that BTH would always act as a functional analog of SA and would also activate the SAR response in the same manner as pathogen infection.

In one example, global gene expression changes in *Arabidopsis* are monitored under different SAR inducing or SAR-repressing conditions using a DNA microarray with 10,000 ESTs. 660 ESTs differentially expressed under two or more of these conditions are chosen to derive groups of genes with common regulation patterns, or regulons. The regulon containing the PR1 gene, a reliable marker for SAR induction in *Arabidopsis,* contains other PR-genes and also defines additional novel genes likely to function during SAR and pathogen response. A common promoter element is identified in genes of the PR1 regulon. This element is bound by members of a plant-specific transcription factor family. Our results confirm expression profiling as a means to both define regulation networks and accelerate pathway-specific gene discovery in plants.

Hence, in one embodiment, the present invention is directed to a method for isolating a regulatory DNA sequence from a differentially expressed gene, comprising:

(a) obtaining an expression profile of a genome under two or more different conditions;

(b) comparing gene induction and repression patterns among each of the different conditions;

(c) identifying a gene that is differentially expressed among the different conditions;

(d) obtaining the sequence of genomic DNA comprising the gene identified as being differentially expressed and regulatory elements associated with said gene; and

(e) isolating a regulatory DNA sequence associated with the differentially expressed gene from the genomic DNA.

In another embodiment, the present invention is directed to a method for isolating a common regulatory DNA sequence from a group of co-regulated genes, comprising:

(a) obtaining an expression profile of a genome under two or more different conditions;

(b) comparing gene induction and repression patterns among each of the different conditions;

(c) identifying a group of genes that are co-regulated with respect to each other but that are differentially expressed with respect to the rest of the genome among the different conditions;

(d) obtaining the sequence of genomic DNA comprising the genes identified as being differentially expressed and regulatory elements associated with said genes;

(e) identifying common regulatory DNA sequences among the genomic DNA associated with the co-regulated genes; and

(f) isolating a common regulatory DNA sequence associated with the differentially expressed genes from the genomic DNA.

In a preferred embodiment, said regulatory DNA sequence is a promoter. In another preferred embodiment, said genome is a plant genome. In yet another preferred embodiment, said expression profile is obtained using a DNA microarray. In one preferred embodiment, said two or more different conditions comprise biotic stimuli. In another preferred embodiment, said two or more different conditions comprise abiotic stimuli. In yet another preferred embodiment, said two or more different conditions comprise biotic and abiotic stimuli.

According to a preferred embodiment, said genome is a plant genome and wherein said two or more different conditions comprise at least one SAR inducing condition or at least one SAR repressing condition. Preferably, said least one SAR inducing condition comprises pathogen infection, SA application, BTH application, *NIM1* gene expression, or a *cim* mutation, and preferably said least one SAR repressing condition comprises NahG expression or a *nim1* mutation.

In an additional embodiment, the present invention is directed to a method for isolating a regulatory DNA sequence from a differentially expressed plant gene, comprising:

(a) using DNA microarray technology to obtain an expression profile of a plant genome under two or more different conditions, wherein said conditions comprise at least one SAR-inducing or SAR-repressing condition;

(b) comparing gene induction and repression patterns among each of the different conditions;

(c) identifying a gene that is differentially expressed among the different conditions;

(d) obtaining the sequence of genomic DNA comprising the gene identified as being differentially expressed and regulatory elements associated with said gene; and

(e) isolating a regulatory DNA sequence associated with the differentially expressed gene from the genomic DNA.

In still another embodiment, the present invention is directed to a method for isolating a common regulatory DNA sequence from a group of co-regulated plant genes, comprising:

(a) using DNA microarray technology to obtain an expression profile of a genome under two or more different conditions, wherein said conditions comprise at least one SAR-inducing or SAR-repressing condition;

(b) comparing gene induction and repression patterns among each of the different conditions;

(c) identifying a group of genes that are co-regulated with respect to each other but that are differentially expressed with respect to the rest of the genome among the different conditions;

(d) obtaining the sequence of genomic DNA comprising the genes identified as being differentially expressed and regulatory elements associated with said genes;

(e) identifying common regulatory DNA sequences among the genomic DNA associated with the co-regulated genes; and

(f) isolating a common regulatory DNA sequence associated with the differentially expressed genes from the genomic DNA.

Preferably, said regulatory DNA sequence is a promoter. Also, preferably, said least one SAR inducing condition comprises pathogen infection, SA application, BTH application, *NIM1* gene expression, or a *cim* mutation, and said least one SAR repressing condition comprises NahG expression or a *nim1* mutation. In an especially preferred embodiment, said two or more different conditions comprise BTH application, and at least one condition selected from the group consisting of pathogen infection, SA application, and a *nim1* mutation. In a most preferred embodiment, said two or more different conditions comprise BTH application, pathogen infection, and SA application, and wherein one or more genes are identified that are inducible by BTH application but not by pathogen infection or SA application. In another especially preferred embodiment, said two or more different

conditions comprise BTH application and a *nim1* mutation, and wherein one or more genes are identified that are inducible by BTH application in *nim1* mutant plants. It is desirable in the above embodiments that said one or more genes are inducible at least 5x by BTH application.

DETAILED DESCRIPTION OF THE INVENTION

Understanding the regulation of plant defense responses is a central goal of current plant molecular biology. Research has mainly focused on the initial recognition of the pathogen by the plant and on the identification of effector molecules that may be responsible for plant defense. This led to the concepts of race-specific resistance and disease resistance marker genes. Many marker genes are classified as pathogenesis related (*PR*) genes that are induced following attack by a wide range of pathogens. A subclass of *PR* genes, termed SAR genes, is induced concomitant with the onset of SAR. The expression of the SAR gene *PR-1* in *Arabidopsis* is directly correlated with the SA content in the plant making *PR-1* expression an appropriate marker for genetic SAR induction. The recent functional and structural analysis of the *Arabidopsis PR-1* promoter revealed several INA-responsive elements, as well as putative *cis* elements not involved in the SAR response (Lebel *et al.*, 1998). Thus *PR-1* transcription may not be strictly limited to defense responses of plants to pathogen attack, and in fact *PR-1* expression has been observed in other circumstances as well (Uknes *et al.*, 1993). The PR-1 peptide is neither required nor sufficient to confer complete disease resistance. Rather, it is thought that the combination of many factors (chitinases, glucanases, antimicrobial peptides) together confer the broad-spectrum disease resistance observed during SAR. Based on structural similarity to small antimicrobial peptides, PR-1 is likely to be one of the factors contributing to this resistance.

From the structural analysis of several other promoter sequences of *PR* genes, which do not necessarily share the same *cis* elements (reviewed by Rushton and Somssich, 1998; Yang *et al.*, 1997), it can be derived that the signal transduction pathway is diverging from a common element in this cascade, salicylic acid. Consequently, a screen based on PR-1 as the sole marker can neither identify all members of the signaling cascade nor be totally specific for mutations in the SAR signaling cascade. But it can identify some signaling elements that

link the pathogen recognition event with the multi-component defense response. The goal of this study was to identify an array of different mutants by monitoring for the expression of one marker gene. These mutants were subsequently analyzed to further our understanding about SAR regulation.

1. Principle of the screen for *cim* mutants

To isolate new disease resistance mutants we screened for plants that constitutively expressed the *PR-1* gene. To monitor *PR-1* gene expression, we chose the luciferase reporter gene because luciferase activity can be monitored *in vivo* without affecting the integrity of the plant. This feature opens up the possibility to rapidly examine many plants as well as to reexamine the same tissue several times throughout the experiment (Millar *et al.*, 1992). Also, luciferase activity can easily be reexamined *in vitro* providing a mean for fast confirmation and quantification of results obtained by *in vivo* monitoring. Two lines of evidence correlate light emission by *PR-1*/luciferase plants with SAR gene expression:

- Treatment with chemical activators of SAR gene expression induces luciferase activity *in vivo* and *in vitro* with kinetics similar to those seen for *PR-1* mRNA levels.
- Infection of *PR-1*/luciferase plants with incompatible pathogens activates luciferase activity and *PR-1* gene expression in both local and systemic tissue.

Taken together, these results strongly suggest that in the selected *PR-1*/luciferase line, *in vivo* monitoring of luciferase activity represents a method for detection of the onset of SAR in these plants (within the biological specificity of *PR-1* gene induction), and that this method can be used to identify mutants that display both constitutive SAR gene expression (*PR-1*, *PR-5*) and enhanced resistance to virulent pathogens. In addition, this reporter line could be used for detailed studies of tissue localization and timing of *PR-1* gene expression, as this has been done before in parsley leaves using *in situ* hybridization (Reinold and Hahlbrock, 1996).

The size of the EMS mutant screen was designed to near-saturate the genome with point mutations that might lead to constitutive *PR-1* gene expression. We estimate that the mutation rate corresponded to a 400-fold increase of the natural mutation frequency (which has been estimated to 5×10^{-4} mutations per gene per generation; Feldmann *et al.*, 1994; Li and Redei, 1969). In previous screens for SAR mutants based directly on biological tests

(Delaney *et al.*, 1995; Glazebrook *et al.*, 1996) or on the use of other reporter genes (Bowling *et al.*, 1994; Cao *et al.*, 1994), the M1 and for the most part the M2 populations were much smaller. The ease of mutagenesis by EMS is countered by the difficult cloning of the underlying (point) mutation, since EMS introduces primarily G-A transitions through O^6 alkylation of guanine (Britt, 1999). Other commonly used mutagens in *Arabidopsis* include foreign insertion sequences, such as T-DNA or heterologous transposons that have been engineered to carry selectable marker genes (for review, see Stiekema and Pereira, 1998). While transposons (for example the *En/Spm* system; Cardon *et al.*, 1993) move preferentially to closely linked locations and consist commonly of a two-component system (transposase and mobile element), T-DNA insertions occur randomly throughout the genome and do not need to be mobilized (Schulz *et al.*, 1994). Although only a few mutations per genome are introduced by T-DNA mutagenesis (thus increasing the required mutant population size), more mutations result in a detectable phenotype than with EMS mutagenesis. Furthermore, T-DNA mutagenesis facilitates the cloning of the mutant gene (especially in cases where the phenotype is difficult to determine). We therefore also tried a *cim* screen with 10,000 T-DNA 'tagged" mutant lines in the *PR-1*/luciferase background. The frequency of *cim* mutations found in the two screens cannot be compared because the second screen was done in the T1 generation.

Characterization of the different mutants validated the approach taken since all mutants with increased luciferase activity exhibited increased resistance to several virulent pathogens. The mutants fall into two classes, *lsd* mutants, the predominant class, and *cim* mutants. Since we were interested in mutants in SAR signal transduction, and not in mutants in which SAR is induced as a consequence of spontaneous cell death, we focused the study on *cim* mutants. *cim* mutants define a diverse group of loci with different disease resistance spectra.

2. Genetics of cim mutants

Interestingly, our EMS screen for *cim* mutants yielded mainly dominant or codominant mutations, which rendered genetic analysis, complementation tests, and pathway classification by epistasis studies more difficult than with recessive mutants because they require additional crosses, more generations and analysis of larger populations. We did, however, map four of

the *cim* mutants, demonstrating that several independent loci have been identified. Several hypotheses may be cited to explain the prevalence of (co-) dominant mutants. The SAR signaling cascade might be regulated by functionally overlapping signal transduction pathways. It has been estimated that 60% of all *A. thaliana* genes are present in multiple copies, based on sequence analysis. Parallel signaling cascades leading to SAR may be favored over single-gene regulation to counter the selection pressure exerted by the pathogen to overcome the plant defense. Thus, only dominant gain-of-function mutations would be phenotypically detectable. Alternatively, SAR signaling might be regulated by a cascade of repressors, in analogy to for example the pathway controlling apoptosis in *Drosophila* and *C. elegans* (McCall and Steller, 1997; Vaux and Strasser, 1996). A complex network of repressors (bcl-2/ced-9, p35, and others) keep the cell alive and the loss of one of these factors is sufficient to induce apoptosis. Similarly, the SAR signaling cascade would be turned on by a loss-of-function mutation in a negative regulator, and this mutation could be dominant or appear to be dominant as a result of haplo-insufficiency (Melnick *et al.*, 1993).

The *cim* mutations are often not expressed in 100% of the self-progeny. Incomplete penetrance has also been found in several hormone mutants (for instance impaired in the ethylene or giberellic acid pathways; Kieber, 1997; Ogas *et al.*, 1997) and also in SAR mutants (e.g. *lsd2*, *lsd7*). To date, no genetic explanation for incomplete penetrance of mutations in isogenic backgrounds has been found. Exogenous and endogenous events, such as heat, light, and cell homeostasis can be decisive in triggering a labile switch, as known in less complex biological systems, such as the phage lambda lysis-lysogeny decision (McAdams and Shapiro, 1995; Ptashne, 1992). This bistable behavior is well known in biological network theory, and occurs also in higher eukaryotic cells: Feedback regulation can result in the presence of two discrete steady-state activities, such that a small stimulus is sufficient for a cell to trigger the transition to one state and to stabilize the cell in this state (Bhalla and Iyengar, 1999). Such a labile equilibrium would be in accordance with haplo-insufficient mutations. The loss of one gene copy of a regulatory factor might still allow proper regulation until stress situations titer this factor and the mutation becomes phenotypically evident.

Interestingly, incomplete penetrance in mixed genetic background is a common feature of many mutants in disease resistance, found not only in *Arabidopsis*, but also in maize. This phenomenon may reflect race-specificity observed in many plant-pathogen interactions. Race-

specificity may not be limited to *R* genes, but may extend to parts of the signaling cascade leading to systemic resistance. It is known that several marker genes like thionins and CXC750 are induced in an ecotype-specific manner (Aufsatz and Grimm, 1994; Epple *et al.*, 1998). The incomplete penetrance in mixed backgrounds rendered positional cloning difficult. Penetrance of the Cim phenotype in mixed ecotype backgrounds was for all tested mutants low (from 5% of the expected ratio to 20% in *cim*713 for which 15% of the F2 outcross progeny displayed the phenotype instead of 75% expected). Positional cloning was further hampered because *cim*713 lies in a genomic region with extremely low recombination and mutation frequency. It is interesting to note that the *NIM1/NPR1* gene is situated in a similar genomic region (Ryals *et al.*, 1997).

There are a number of reasons why the screen did not recover multiple alleles of any genes, despite the thorough genome coverage of the EMS screen. The observed incomplete penetrance is one possible explanation. Another possible explanation is that different mutations in the same gene may cause either lesion mimic or Cim phenotypes.

3. Nature of *cim* mutants

Ten *cim* mutants were further characterized. While all these mutants were free of HR-like lesions, leaf and floral morphologies in some mutants deviated from wild-type. Two mutants had a wild-type morphology, suggesting that the morphological changes in the other *cim* mutants were due to mutations independent of the SAR mutation. However, three backcrosses did not separate the morphological changes from the SAR inducing mutation. Furthermore, the two wild-type looking mutants (*cim*328 and *cim*658) fell in the class of the weakest *cim* mutants, based on SA accumulation and *PR-1* gene expression. It is therefore impossible at this point to distinguish between a causally related morphological change (i.e. SA as a morphogen) or an organ malformation induced as a consequence of the physiological cost of constitutive SAR induction (potential metabolic sinks include *PR* gene expression and SA accumulation). In experiments with repeated applications of chemical SAR inducers, no physiological cost of SAR induction was identified. Since these trials were performed on mature plants, they do not account for possible nutrition shortage during early development. However, metabolic engineering in plants has often shown that plants can produce huge

amounts of secondary metabolites or storage polymers without showing energy deficiencies. For instance, ectopic expression of a tryptophan decarboxylase in tobacco led to the accumulation of 1 milligram tryptamine per gram fresh weight of plants with unchanged morphology (Sonstad *et al.*, 1990). This value can be compared to up to 4 milligram salicylic acid per gram fresh weight in the *cim* mutants, which corresponds to a 15-fold increase to uninduced wild-type levels. It should be noted that ectopic expression of tryptophan decarboxylase in potato resulted in a severe alteration of the phenylpropanoid pathway. The plants were morphologically unchanged, though hypersusceptible to pathogens because of depletion of the chorismate pool (Yao *et al.*, 1995). Plant metabolism appears to be very flexible in reacting to exogenous manipulations, such as the inhibition of amino acid biosynthesis (Guyer *et al.*, 1995), or the perturbation of the carbohydrate homeostasis (Stitt *et al.*, 1990). Plants that overexpressed a yeast invertase increased the glycolysis and were able to define new "Sink" and "Source" tissues (Sonnewald *et al.*, 1991). Similarly, it is likely that *cim* mutants can compensate for the increased energy flux through the SAR pathway by enhancing and deviating the metabolism. *cim* mutants often appear darker green, and might have a higher chlorophyll content than wild-type. It is also remarkable in this context that *cim* mutants show enhanced transcription of genes encoding functions in energy metabolism, photosynthesis and protein biosynthesis. Similar results in parsely cells also revealed extensive changes in metabolism during fungal infection (Batz *et al.*, 1998).

Using differential resistance against pathogens, we can assemble a first-order classification of the *cim* mutants. Several mutants were highly resistant to all tested pathogens (e.g. *cim*677, *cim*810). Others were only resistant to fungal pathogens (e.g. *cim*205). Mutant *cim*695 was resistant to *Pseudomonas* and *Peronospora* spp., but not to *Erysiphe*. Mutant *cim*328 was strongly resistant only to *Erysiphe*. The identification of an *edr*-like mutant such as *cim*328 with weak accumulation of *PR* gene transcripts was possible because the luciferase reporter gene assay is very sensitive.

It will be interesting to check for resistance to other pathogens, including viral, and root-specific (*Pythium*, nematodes) pathogens. These tests, as well as tests with non-SAR pathogens (e. g. *Fusarium*), are in progress, and will help to further define regulation and branch points in the induced plant disease resistance response.

As all *cim* mutants are resistant to several virulent pathogens, the resistance appears *R* gene-independent and does not require an HR. However, some *cim* mutants (for example *cim*695, *cim*716) develop a strong HR-like lesion formation in response to *Peronospora parasitica* inoculation. Most *cim* mutants are able to develop an HR in response to an avirulent bacterial pathogen but some appear to simply bypass HR (*cim*713, *cim*810). *Arabidopsis* mutants with a similar phenotype, called *dnd* (defense, no death) have recently been isolated (Yu *et al.*, 1998). They exhibit disease resistance to two virulent pathogens and do not develop an HR when inoculated with avirulent *Pseudomonas* strains, although they remain fully resistant. These *cim* and *dnd* mutants raise questions about the role of cell death in defense. Obviously, cell death is not required for an effective, incompatible gene-for-gene interaction, or for the induction of SAR in these mutants. HR may be required in wild-type to potentiate an SAR-inducing signal, possibly via the release of reactive oxygen species (Alvarez *et al.*, 1998).

It is tempting to speculate about the mechanistic nature of broad-spectrum disease resistance in the *cim* mutants. Because of the very different lifestyles of the pathogens used (*Erysiphe*, *Peronospora* spp., *Pseudomonas*), it is unlikely that simple host morphological changes, for example in the cuticle, are responsible for this resistance. Rather, the *cim* mutations cause a fundamental deregulation of the SAR pathway. In principle, hypermorphic gain-of-function mutations or haplo-insufficiencies may be able to activate the SAR pathway through cross-talk and side effects without being members of the signaling pathway *per se*. This might explain the different disease resistance spectra of some mutants. However, the same two classes of mutation could exist within the *R* gene-dependent signal cascades, although the resistance in *cim* mutants is not race-specific. A preinduction of the *R* gene-mediated signaling pathway could induce the cascade leading to an activation of SAR, and consequently convert a compatible interaction into an incompatible. In principle, even a mutation in an *R* gene could trigger SAR. It has previously been shown that such mutations can also cause a lesion mimic phenotype (shown for *Rp1* in maize (Hu *et al.*, 1996), and *mlo* in barley (Peterhänsel *et al.*, 1997)).

How far beyond SA and *PR* transcript accumulation can the comparison between *cim* mutants ("hypochondriacs") and pathogen-infected plants be stretched? Upon infection by pathogens, plants undergo fundamental cellular and biochemical changes (Gus-Mayer *et al.*,

1998). *Arabidopsis* is not easily amenable to study morphological alterations on cellular level, such as cytoskeletal rearrangements, or increased ion fluxes through the plasma membrane. Measurable parameters include gene expression (*PR-1*), accumulation of secondary metabolites (SA, camalexin) and apposition of cell wall material (e.g. callose). Parallels of the measurable biochemical changes in the *cim* mutants to pathogen-infected plants are restricted to systemic changes induced by pathogens and do not reflect the plant's local defense responses.

The quantitative differences in disease resistance and biochemical markers among the *cim* mutants reveal a complex regulation pattern of the different signaling branches of disease resistance responses in *Arabidopsis*. To assess the phenotypic diversity of the *cim* mutants at molecular level, we utilized a highly paralleled gene expression profiling method, the DNA microarray. We monitored gene expression of roughly 5000 different genes in the *cim* mutants and following chemical induction of SAR. While expression of approximately one quarter of all *Arabidopsis* genes was monitored, pathogen-induced genes are likely to be underrepresented in the MSU *est* collection because the cDNA libraries were not derived from pathogen-induced tissue. The deciphering of the *Arabidopsis* genome will allow the monitoring of expression of all genes, as has been done for *S. cerevisiae* (DeRisi *et al.*, 1997). In the meantime, *ests* are the best way to access genes in *Arabidopsis*. We showed that DNA microarray experiments can be compared to each other, using internal standardization. Internal control RNA also allowed us to estimate the sensitivity of detection. 0.06 ng spiked RNA was detected (at 1000 FSI just above background) which corresponds to a copy number of 1 per cell (based on 100,000 mRNA species per cell). The distribution of abundancies of mRNA species obtained by the microarray matched well the distribution curve obtained by other methods such as R_{ot} kinetics (Kamalay and Goldberg, 1980). In addition, in comparisons to Northern blot analysis, we found very similar sensitivities for the DNA microarray in detecting differences in gene expression. The similar sensitivity of Northern blot analysis and DNA microarray is based on similar hybridization stringencies: The washing conditions of Northern blots (65°C, G + C buffer at 71 mM $[Na^+]$, DNA-RNA interactions) and DNA microarrays (30°C, 0.2 x SSC at 39 mM $[Na^+]$, DNA-DNA interaction) have quite similar stringencies (but the estimates are hampered by the lack of exact formulas). Empirically, it can be assumed that cross-hybridization on the DNA microarray does not occur

to a greater extent than during Northern blot analysis, although cross-hybridization has to be taken into account for closely related gene family members with both methods (e.g. LTPs). The DNA microarray cannot distinguish between related transcripts of different size, whether from cross-hybridizing family members of from differential splicing (e.g. XET). Both sensitivity and reproducibility of microarrays are not yet matched by the other methods we used to assess the molecular differences in SAR induction. Two-dimensional protein gel electrophoresis suffers from insufficient resolution and difficulty in peptide identification. Metabolite profiling depends largely on the extraction protocol and on the detection method and hence detects only subsets of metabolites, for example compartmental or structural. Besides these technical limitations, it is currently unknown how many changes in secondary metabolites are induced during plant pathogen defense. In a similar study, in barley, only a few changes were detected in the biochemical cytosolic and cell wall composition during pathogen infection (von Roepenack *et al.*, 1998). In conclusion, it is not straightforward at this time to match proteins to gene induction, and metabolites to proteins. Hence it is difficult to match traits to genes.

To determine if correlations could be detected between pathogen-specificities and gene induction patterns, we chose from our collection three *cim* mutants that displayed different spectra of disease resistance. Despite clear morphological differences among the three *cim* mutants and the BTH control, many genes were altered in similar ways in the four experiments, compared to wild-type. Co-regulated sets of genes could be used as more reliable indicators of specific induced defense responses than single marker genes. Included in this group are the known SAR marker genes *PR-1*, *PR-2*, and *PR-5*, other pathogen-inducible genes such as extensins (Memelink *et al.*, 1993) and CXC750 (Aufsatz and Grimm, 1994), as well as genes encoding proteins with unknown function.

A breakdown of the 478 elements with altered expression in the *cim* mutants into gene groups based on presumed function, shows that the largest class contains genes with no homologies or unclear functions (39%, table below). In the *Arabidopsis* genome (represented here by the ESSA1 contig, 1.9 Mb; Bevan and al., 1998) this number is higher, probably because many of the unknown genes are not highly transcribed and therefore not in the *est* collection. Results obtained from the *C. elegans* sequencing project support this hypothesis: conserved genes are more likely to be highly expressed (The *C. elegans* sequencing

consortium, 1998). Noteworthy is the high percentage of induced genes involved in disease and defense. 14% of the genes with altered expression in the *cim* mutants fall into this class compared to 7% in the ESSA contig. This is remarkable because the ESSA contig contains an MRC and therefore more genes from this class than the genome average (estimated at 2 - 3%; Botella *et al.*, 1997) and because in the *est* population, pathogen-induced *ests* are most likely underrepresented. For instance, no *est* corresponding to the *PR-1* gene is present in the MSU collection.

Classification of genes with altered expression in the *cim* mutants and comparison to prevalence of these gene classes in the genome (as represented by the ESSA1 contig).

gene class	gene expression (%)	genome (ESSA contig) (%)
total number	100 (n = 478)	100 (n = 485)
metabolism	9.8	10.4
energy/ photosynthesis	4.7	2.9
cell growth and division	1.6	2.9
transcription	5.0	7.8
protein synthesis	4.5	1.6
protein transport and storage	3.4	2.6
transporters	1.6	0.8
intracellular traffic	1.3	1.0
cell structure	6.1	3.6
signal transduction	6.6	4.4
disease/defense	14.0	7.3
secondary metabolism	1.3	4.7
transposons	0.5	2.6
unclear	22.5	31.7
no homology	16.9	15.8

Genes for protein synthesis and energy production are induced disproportionally in the *cim* mutants. Genes for secondary metabolism, induced only under specific conditions, and transposons, usually not transcribed, are underrepresented in the collection of induced genes, compared to the percentage of genes in the genome. Unfortunately, no classifications for *est* libraries similar to the one shown in the above table have been published. Thus, no comparisons can be made between the numbers of genes of various classes in the *est* population in general versus those with altered expression in our mutants, nor can any statements be made regarding the representation of the genome in the *est* collection.

The prevalence of common features in gene expression among the mutants lead us to the hypothesis that pathogen specificity in (genetically) induced SAR originates in subtle

differences of gene expression rather than the presence of several independent signaling pathways represented by distinct sets of marker genes. The high degree of similarity can be better expressed by statistical analysis than by gene expression profiles. We applied statistical methods normally used to calculate phylogenetic distance among macromolecules to calculate the similarities of gene expression patterns in the *cim* mutants and their crosses to NahG. The gene expression patterns in three mutants are indeed very similar, although distinct from each other. Mutant *cim*205 best mimics the chemical SAR induction by BTH. This is for example reflected in the suppression of the senescence associated proteins in these two samples that is not found in the other two mutants.

Furthermore, we addressed the question of what genes were regulated in a similar way under various conditions. To form these groups of "regulons" (groups of co-regulated genes), we used the same calculation methods to find similarities in gene expression patterns under 16 different conditions, including the 6 different experiments listed in table 9, and 10 others (including herbicide treatments and pathogen treatments). The *PR-1* gene is almost unique in its expression regulation, clustering only with few other genes that match its expression pattern under these 16 different conditions. The three copies of *PR-5* and of the *PRXC* gene, as well as 20 other *ests,* are included in this cluster.

In addition to the identification of common patterns, we were also able to correlate specific gene groups to specific resistance phenotypes, but the causality of these correlations remains unclear for now. Causalities have in the past been proven to be difficult to establish from correlations. For instance, the role of the PR-1 protein in plant pathogen defense is still not clear and the function of the phytoalexin camalexin is only starting to be elucidated (Glazebrook *et al.*, 1997). Genetic studies addressing this question are complicated by functional overlap of the defense-related molecules and by the fact that quantitative differences in gene induction seem to be more important than qualitative differences in generating pathogen resistance specificities, as derived from our mutant characterization. Using all the *cim* mutants in our collection in DNA microarray experiments, it may be possible to match specific gene inductions to specific resistance spectra.

By counting the total number of genes with altered expression, we concluded that *cim*713 was likely to display a more complete resistance than the other two mutants, and this was confirmed by pathogen tests. It is, however, not easy to distinguish the influence of weak

and strong alleles from hierarchical gene induction effects. Finally, it has to be stressed that the regulation of gene transcription represents only one (important) level of control of cell activity. In higher eukaryotes, significant regulation also occurs at the translational and protein levels (hence the importance of 2D gel electrophoresis). In spite of its limitations, it can be anticipated that expression profiles will provide insight into gene expression regulation, of how, when, where and why the genes act in concert to regulate complex processes, such as pathogen defense, in a whole organism (Ruan *et al.*, 1998).

4. Epistasis studies and the relationship of *cim* mutants to salicylic acid

Along with mutant classifications based on disease resistance spectra and gene expression profiles, epistasis studies to known disease resistance regulatory mutants, such as *ndr1* (Century *et al.*, 1995), *eds1* (Parker *et al.*, 1996), and *pad4* (Glazebrook *et al.*, 1997) will give further insight into the relative relationships among the loci identified by this collection of *cim* mutants. Because of genetic problems encountered during the work with the *cim* mutants, we limited the epistasis analysis to crosses of *cim* mutants to the dominant NahG plants, asking the question whether SA accumulation is a prerequisite of SAR induction in the *cim* mutants. From the study of several *lsd* mutants in the NahG background, it has been concluded that SA accumulation is regulated by a feedback regulation (Weymann, 1995). Similarly, biochemical studies have shown that PAL, a key regulator of the SA biosynthesis, is itself activated by SA (Mauch-Mani and Slusarenko, 1996). As a consequence, a mutation "downstream" of SA accumulation, such as a dominant-negative mutation in the *NIM1/NPR1* gene might lead to both a Cim phenotype, and to SA accumulation by feedback activation. This could explain why no *cim* mutants without elevated levels of SA were identified. Mutants of this class might still display disease resistance when the SA accumulation is depleted by an SA hydroxylase (NahG).

We in fact found two mutants *cim*695 and *cim*713 that exhibited SA-independent resistance to *Peronospora*, although *PR-1*/gene expression in all *cim* mutants was strictly dependent on SA accumulation. A similar phenotype as the one observed in these two *cim* mutants has been described for *cpr5* and *cpr6*, which retained resistance to *P. parasitica* in a *nim1/npr1* background while resistance to virulent *Pseudomonas syringae* strains was

suppressed (Bowling *et al.*, 1997; Clarke *et al.*, 1998). Disease resistance of these mutants in a NahG background, however, has not been reported. While *cpr5* in a NahG background exhibits an extreme dwarf phenotype with black lesions (Bowling *et al.*, 1997), mutants *cim*695 and *cim*713 retained wild-type morphology in the NahG background. The common induction of SA-independent and SA-dependent disease resistance pathways in *cim*695 and *cim*713 might illustrate that their signaling cascades share common members downstream or independent of cell death.

Furthermore, *cim*695 and *cim*713, which exhibit an SA-independent resistance, do accumulate SA to five fold higher levels than wild type. The resistance conferred by mutations *cim*695 and *cim*713 may therefore lie in a feedback loop as suggested above, triggering multiple resistance mechanisms including SA-independent resistance pathways which lead subsequently to SA accumulation. SA-independent resistance has been described in the literature. For instance, a jasmonate-dependent defense response in *Arabidopsis* has been shown to confer resistance to a distinct set of pathogens (Thomma *et al.*, 1998). This wound- and necrotrophic-inducible disease resistance is correlated with the expression of the *PDF1.2* gene (Penninckx *et al.*, 1998).

We used the collection of *cim* mutants to address whether *PDF1.2* gene expression is correlated to SA-independent disease resistance, as suggested for other broad-spectrum disease resistance mutants (Bowling *et al.*, 1997). We found mutants that, when crossed into a NahG background, expressed the *PDF1.2* gene at high levels, but lost resistance to *P. parasitica* (e.g. *cim*205, *cim*716). Conversely, *cim*713 in a NahG background retained resistance to *P. parasitica* despite a severe reduction in *PDF1.2* expression. Thus, in our mutant collection, there is no consistent correlation between the transcription of the *PDF1.2* gene and the activation of an SA-independent signaling pathway that confers resistance to *P. parasitica*. This may imply the presence of a third, SA-independent pathway leading to resistance to *P. parasitica Noco2*, which does not include induction of *PDF1.2*. Induced systemic resistance (ISR) has been found to be both independent of SA and not linked to *PDF1.2* expression (Pieterse *et al.*, 1998). Because no molecular marker or mutant defines ISR to date, it is not possible to compare the SA-independent resistance observed in the *cim* mutants to this phenomenon.

The difference in resistance of the "double mutants" is reflected in a similarity tree of gene induction in the mutants. The radiation between *cim*713 and *cim*205 in NahG backgrounds is bigger than it is in the mutants alone. But both resemble NahG plants more than they resemble the mutants. This addresses the question concerning the role of SA in plants. More than 300 genes show altered expression in NahG plants compared to wild-type. The removal of SA (and/or the accumulation of catechol) triggers a major switch in gene regulation, as does the accumulation of SA. Using the new gene expression profiling technology, we were able to visualize this global gene switch. There might be a well-defined master gene switch (such as a homeobox regulation (Gehring, 1987), see Plesch *et al.*, 1997 for an example in *Arabidopsis*) that is naturally induced by salicylic acid accumulation, and that can also be triggered by exogenous BTH. Based on this and the pleiotropic effects on plant development caused by increased SA accumulation (reminiscent of mutants impaired in hormone action; Kauschmann *et al.*, 1996), the role of SA as a plant hormone may need to be revisited. SA, as ethylene and jasmonic acid, might have more functions than the induction of disease resistance. The regulation of all three signal molecules is interdependent and the role of the two latter molecules in the *cim* mutants remains to be elucidated.

5. The Transcriptome of *Arabidopsis* During SAR

To gain a more thorough understanding of the complex regulation of gene expression during SAR, and to identify new (marker) genes associated with SAR, we used expression profiling (Schena *et al.*, 1995; Shalon *et al.*, 1996). PolyA+ RNAs obtained from plants grown under 18 different conditions that either induce or repress SAR (see table below) were used in mixed hybridizations against RNA of untreated wild-type plants on a DNA microarray (Ruan, *et al.*, 1998) with 10,000 ESTs, a Unigene set representing roughly one third of all *Arabidopsis genes*. In addition, two treatments which alter plant metabolism, but are not related to SAR, were analyzed (see table below). Some samples were taken during the induction phase of SAR (4 hours after BTH induction), or from primary, pathogen-infected tissue, but most of the samples were derived from (steady state) SAR maintenance phase (e.g. 48 hours after BTH induction, cim mutants). Together, the experiments gave us 1.8×10^5 gene expression data points. Under these 18 conditions, however, only 660 different genes

displayed significant differential expression compared to wild-type in at least two SAR-relevant samples (significance level defined as 2.5 fold differences from wild-type). By comparing the fluorescence signal intensities to spiked controls, abundance of mRNA species can be determined . Most of the genes with altered expression fall into the class of low-abundance transcripts (1 to 10 copies per cell), while a few were in the medium abundance class. Housekeeping genes, with more than 100 copies per cell, did not exhibit altered gene expression under SAR.

Diversity of conditions used to describe the transcriptome of *Arabidopsis thaliana* during SAR.

Sample (all compared to wild-type tissue)	Comments
Adenylosuccinate synthetase antisense Hydantocidin (0.9 mM)	Reference inductions, irrelevant to resistance
cim6 *cim7* *cim11*	Constitutive SAR mutants
NIM1 overexpresser	Primed SAR response
NahG *cim6NahG* *cim11*NahG NahG + *Pst* DC3000 *avrrpm1* 2' tissue, 44 hours	No SA accumulation, no SAR response
NahG + *Pst* DC3000 *avrrpm1* 1' tissue, 44 hours BTH 4 hours BTH 48 hours	Primary (partial) LAR induction SAR induction SAR maintenance
P.p. EMWA1, 48 hours *Pst* DC3000 *avrrpm1* 1' tissue, 44 hours *Pst* DC3000 *avrrpm1* 2' tissue, 44 hours *nim* + Pst DC3000 *avrrpm1* 2' tissue, 44 hours *P.p.* Noco5, 48 hours	Compatible (disease causing) interaction Incompatible interaction

First, we compared gene induction patterns of these 660 genes among each of the different conditions. Data representing a series of differential gene expression measurements was obtained. Analysis of both the regulation of individual genes under varying conditions or of the overall similarity of the various conditions based on expression profiles required the calculation of distance matrices. These were calculated in S-Plus using the dist function under the euclidean metric. Clustering of similarly regulated genes was performed and visualized in S-Plus using hierarchical clustering under the compact method with the, functions hclust and plclust. Trees used to depict similarity in gene expression patterns among the conditions may

be produced using the fitch and drawtree programs distributed in the Phylip suite (Phylogeny Inference Package) version 3.57c.

Genetic, chemical and biological induction of SAR all caused very similar patterns of gene induction. The three *cim* mutants, which constitutively express SAR, have a pattern of transcriptional induction similar to that caused by BTH during SAR maintenance. Interestingly, the overexpression of *NIM,* which primes the SAR response, also resembles SAR maintenance (U.S. Patent No. 6,031,153; Cao *et al.*, 1998; Maleck *et al.*, 1998). Compatible and incompatible races of the pathogens *Peronospora parasitica and Pseudomonas syringae* induced many common genes, although the pattern of gene induction was not identical to either of the BTH-treated samples. In contrast to the requirement for SA accumulation to high levels for SAR induction, many gene inductions are triggered by low levels of SA. For example, in NahG-expressing plants that accumulate only low levels of SA, the expression of more than 300 genes is significantly altered. This emphasizes a general regulatory role for SA in the plant cell.

NahG suppresses SAR gene expression in crosses to two of the SAR-constitutive *cim* mutants, *cim6 and cim11,* to a baseline resembling that of NahG-expressing plants. In addition, NahG expression results in a characteristic gene expression fingerprint in secondary tissue from plants inoculated in primary tissue with avirulent bacteria. This corresponds to the inability of these plants to establish SAR. Interestingly, the corresponding primary tissues in NahG-expressing plants display changes in gene expression which compares very closely to wild-type primary, infected tissue and this sample does not cluster with other NahG samples. This is consistent with findings from grafting experiments that SA is required for SAR development in systemic tissue, but that the systemic signal can be emitted from primary infected NahG-expressing tissue (Delaney *et al.*, 1994; Gaffney *et al.*, 1993; Vernooij *et al.*, 1994).

We used these expression profiles to identify classes of genes that were consistently coregulated, and we sought to identify genes associated with SAR. We used statistical phylogenetic cluster analysis to compare the gene expression profiles among 660 genes under all 18 conditions. Clusters of genes with similar induction patterns under all conditions were grouped. The two clusters most distinct from the PR1 gene induction pattern, with EST 127P22T7 and EST 163B24T7 as "type genes," contain genes that are strongly (more than 20

fold) suppressed by NahG expression. The cluster containing chalcone synthase (ChS) as the type gene includes genes induced during SAR-maintenance, but not during SAR-induction; these are weakly repressed by NahG expression. The cluster containing EST 209E19T7 defines genes that are transcriptionally induced in NahG-expressing plants. The cluster containing EST 118P18T7 defines genes that are not significantly responsive to SAR-inducing conditions like chemical and genetic induction, but do respond to avirulent bacteria and are downregulated in NahG expressing plants. Phenylalanine ammonia lyase (PAL) and 20 other ESTs that cluster together are repressed by NahG expression, but are induced during the maintenance phase of SAR, for example in *cim* mutants or 48 hours after BTH treatment. The cluster of "PR1 like" genes exhibits similar induction behavior to genes in, the PAL gene cluster but these genes are only weakly suppressed in NahG-expressing plants.

To address the question of expression induction relatedness among genes within one cluster, we analyzed the PR1 cluster in detail. The PR1 regulon contained 25 other ESTs (17 different genes). These are prime candidates for SAR marker genes and the encoded proteins are likely to play a physiological role in SAR. The estimated 1.5 to 2-fold redundancy of our EST set is a good internal control for this analysis and we also included three replicates of the PR5 and the PerC cDNAs (as well as 28 other relevant cDNAs) on the DNA-microarray. All three copies of the two genes cluster with PR1, showing the robustness of the DNA microarray analysis. Similarly, two ESTs that are derived from Asparagine synthetase, two from the gene for blue copper binding protein and two from glutathionine S-transferase (GST) cluster with PR1. The standard deviations for differences in expression of the 30 members of the PR1 cluster were smaller than one unit of differential gene expression (compared to 2.5 units cut-off for our definition of significant alterations) under all 18 conditions. Thus, genes in one cluster show highly reproducible expression in our analysis.

To help explain the molecular basis of this co-regulation, we analyzed the promoters of those eight PR1 cluster ESTs encoded within sequenced genomic regions. Strikingly, these genes all shared a common TTGAC consensus cis-element known to bind proteins of the WRKY zinc finger transcription factor binding family, often in multiple copies. In control promoters of eight ESTs selected randomly among the 634 non-PR1-like ESTs, the element TTGAC was only found in the statistically expected frequency. The WRKY element was not present in three of the randomly selected promoters, in three promoters once, in one promoter

twice, and in one promoter three times. WRKY elements are sufficient for defense gene transcriptional induction in different plant systems (Lebel *et al.*, 1998). A functional analysis of the BTH-induction of the PR1 promoter revealed two elements sharing all or four out of five consensus WRKY sites at -676 and - 638 (Lebel *et al.*, 1998). The former acts as a repressor and the latter as a SA responsive element. It is therefore likely that WRKY transcription factors are responsible for the common induction pattern of the eight genes listed in Table 1, and possibly also for the remaining 22 ESTs of the PR1 gene cluster for which no genomic sequence is presently available.

We did not identify another common consensus sequence in the eight PR1 cluster promoters analyzed. The NFκB binding site found in the PR1 promoter which is required for full BTH-induction is not conserved among the eight promoters within the PR1 regulon.

In addition to identification of common regulatory elements, cluster analysis of expression profiles provides a tool to derive physiological functions of genes. This is important for sequences with no close homologs in the databank (for example EST 134C2OT7 or EST 192 K7T7) and also for genes with structural similarity to genes with known function (such as asparagine synthetase). We anchored the results of our statistical analysis first to those genes that were known to be co-regulated with PR1. Interestingly, several genes in the PR1 cluster encode proteins involved in redox regulation. Although reactive oxygen intermediates do not likely act as secondary messengers in SAR downstream of SA accumulation, they are clearly implicated in defense response, either as local second messenger molecules, as direct cellular and microbial toxins or in cell-wall cross-linking and other oxidative processes (Alvarez *et al.*, 1998). From the several known *Arabidopsis* peroxidases, peroxidase C seems to be specifically induced under plant defense conditions, as are some GSTs. Supporting this is the previous demonstration that both peroxidase C and GST are transcriptionally activated in a superoxide-dependent manner in the *Arabidopsis* mutant *Isd1*, which lacks the ability to halt hypersensitive-response-like lesions (Jabs *et al.*, 1996; Dietrich *et al.*, 1994). Likewise, the blue-copper protein might be involved in the regulation of redox stages during SAR or might be the chelator of free cellular copper used for Cu/Zn SOD. Cu/Zn SOD is suppressed during SAR (Fodor *et al.*, 1997).

It has been proposed that plant metabolism adapts to pathogen attack (Batz *et al.*, 1998). We can now more precisely identify such changes and investigate their importance in disease

resistance. For example, it is instructive to speculate about the possible role of the coregulation of glutamine-dependent asparagine synthetase with the PR1 cluster during SAR. An increased flux of carbon through the phenylpropanoid pathway (lower overall N:C ratio) might liberate nitrogen, thus creating a demand for asparagine synthetase as a detoxifying enzyme. Similarly, pathogen infection increases the level of free ammonia in the plant, following an increased metabolism of proteins (Batz *et al.,* 1998). Asparagine synthetase can also utilize ammonia as a substrate. Hence, our observation that asparagine synthetase is co-regulated with PR1 focuses attention on physiological changes during SAR that might not be directly linked to defense. Alternatively, the induction of asparagine synthetase might be related to the induction of tyrosine transaminase. This is also the first time that the production of toxic mustard oils, released by myrosinase from glucosinolates, can be correlated to SAR in *Arabidopsis.*

Our analysis describes the first map of the plant defense transcriptome during SAR in *Arabidopsis.* Using cluster analysis of DNA microarray data, significant changes in gene induction can be differentiated from random correlations. The large number of data points filters out coincidental co-induction and enhances the resolution and significance of serial analysis of expression profiles. It is clear that the description of complex signaling networks can only be obtained by investigating multiple inductive conditions, and not, as before, by pair-wise comparisons. Thus, similar analysis of additional mutants and inducing conditions will further enlarge the complexity and refine the. resolution of the entire *Arabidopsis* transcriptome.

6. Conclusion

We identified novel mutants that mimic the maintenance phase of SAR. Biochemical and molecular analysis revealed a high degree of similarities between these mutants and chemically SAR-activated plants. The mutants differed however in gene expression regulation, metabolite accumulation, genetic map-position and disease resistance spectrum. For instance, an SA- and *PDF1.2*-independent pathway leading to resistance to *P. parasitica* has been identified in one mutant, but not in others. Although questions relating SAR kinetics (induction phase) cannot be addressed in these mutants, the "frozen" steady-state of SAR (or

LAR) presents several experimental advantages that were exploited using three approaches to deepen the understanding of SAR:

• *Epistasis studies*

Epistasis studies with *cim* mutants might help to understand interactions between the different induced disease resistance pathways. The epistasis studies to NahG revealed the presence of SA-independent disease resistance in two *cim* mutants. It will be interesting to establish the spectrum of pathogens affected by this novel resistance. Furthermore, the varying expression of the *PDF1.2* gene in the *cim* mutants raises the question whether the wound-inducible or the ISR pathway are induced or repressed in different *cim* mutants. Mutants impaired in JA- or ethylene-signaling can be crossed to the *cim* mutants that exhibit an SA-independent disease resistance. These double mutants may provide insights into the role of these other signaling molecules in the Cim phenotypes and broad-spectrum disease resistance. Similarly, the direct biochemical measurement of JA and ethylene *in planta*, and a more thorough study of JA and ethylene marker genes in different mutants may provide clues about pathway cross-talk.

• *Gene expression profiling of SAR*

We established a new technology for the study of plant signal transduction, the DNA microarray. Gene expression profiling enabled us to establish correlations between the expression of gene groups and SAR phenotypes (*cim* mutants, NahG). The correlations can in the future be refined with more mutants and with the use of more pathogens to better describe the differential disease resistance in these mutants. It became evident that gene profiling is potentially a powerful tool in understanding metabolic changes, but that it cannot yet (and to some extent will probably never) be able to describe how cells (or even organisms) work. Technology that needs to be improved includes analysis tools to optimize the data capture and mining. For this purpose, we applied general statistical methods to reduce the data complexity to manageable summaries. We were able to select co-regulated genes, "regulons" (which allow the identification of consensus promoter elements), but more sophisticated analysis tools and a higher number of available expression profiles may indeed enable us to visualize the transcriptome of a plant or plant tissue under specific conditions.

The newly obtained gene groups may be useful in developing more precise screens for mutants with specific traits and in lead optimization for agrochemicals. In addition, they may

be helpful in predicting disease resistance spectra in other mutants. Furthermore, gene expression profiles may also be helpful in deriving gene functions, because it provides information on gene expression of many genes under many conditions. In conclusion, gene profiling is a powerful tool in understanding signaling cascades, and their interactions.

Other questions pertaining to the cell metabolism and its regulation require however the use of other techniques. We utilized again the genetic variations in the SAR signaling cascade to follow changes in the proteome and the "metabolome" (entirety of the plant's metabolic products) under SAR induction. However, the observed changes at transcriptional level were not reflected at protein or metabolite level, as detected by 2D gel electrophoresis or HPLC separation. Future technical advances will be helpful to study the metabolic changes during SAR in a systematic way and may also allow the analysis of protein-protein interactions through the large-scale isolation of protein complexes in SAR signaling (Blackstock and Weir, 1999).

• *Gene cloning*

During SAR, a major change in gene expression is induced, that is regulated by a complex network of genetic switches. Several of these switches have been mutated in the *cim* mutants and the genetic cloning will provide a valuable step in understanding plant disease resistance. Several candidate genes have been cloned during this work that await genetic confirmation.

One central regulator gene of the SAR signaling cascade, *NIM1/NPR1* was cloned independently by two groups and seems to be part of a signal transduction cascade with homology to the mammalian IκB/NF-κB pathway (Ryals *et al.*, 1997; Cao *et al.*, 1997; Baeuerle and Baltimore, 1988). Interesting parallels to this conserved pathway in the innate immune response of *Drosophila* and mammals to the defense response of plants have been drawn (Belvin and Anderson, 1996). Besides the homology of several *R* genes to Toll/interleukin-1 receptors (Whitham *et al.*, 1994), members of the kinase cascade may also be shared (*PTO*, *FEN*; Martin *et al.*, 1994; Salmeron *et al.*, 1996). Finally, small antimicrobial peptides, such as defensins, are well conserved among plants and insects (Hancock *et al.*, 1995). Knowing more elements in the cascade will not only allow us to compare this wide-spread signal transduction cascade but also to study molecular interactions,

for instance using the two-hybrid system, and to study physiological effects of overexpression and repression of the genes.

The applications of *cim* genes in agriculture are potentially very promising: Several genes involved in the regulation of various induced defense responses in plants have already been successfully used to engineer disease resistant plants, including the *NIM1/NPR1* gene and the *NDR1* gene in *Arabidopsis* (Cao *et al.*, 1998).

EXPERIMENTAL - MATERIALS AND METHODS

1. Chemicals, Enzymes and Radioisotopes:

All chemicals and organic solvents are purchased, if not otherwise stated, from Sigma Chemicals (St. Louis, MO), Fisher Chemicals (Pittsburgh, PA), or BioRad (Hercules, CA). Chemicals for culture media are obtained from Difco Laboratories (Detroit, MI) or from GibcoBRL Lifescience (Gaithersburg, MD). Luciferin (potassium salt) is purchased from Biosynth (Staad, Switzerland), the detergent SilWet L77 is from Union Carbide Chemicals (Danbury, CT).

Restriction enzymes, T4 DNA polymerase and T4 DNA ligase are purchased from New England Biolabs (Boston, MA) or, if not available from this provider, from Boehringer Mannheim (Indianapolis, IN) or Stratagen (La Jolla, CA). Lysozyme, bovine serum albumin fraction IV and V, and ribonuclease A are delivered by Sigma. For standard PCR, AmpliTaq Gold® from Perkin-Elmer (Foster City, CA) or PCR beads ("Ready-to-go™"; Pharmacia Biotech Inc., Piscataway, NJ) are used. Long range PCR is performed using a special enzyme mix, r*Tth* DNA polymerase (Perkin-Elmer, XL-PCR kit). As DNA size ladder, either the lamba DNA-*Hin*dIII digest (1-23 kb), the phiX174 DNA-*Hae*III digest (0.1 - 1kb), or the 1 kb ladder (1 - 10 kb) from New England Biolabs is employed. Plasmid minipreparations are prepared using Promega's wizard® miniprep kit (Promega Corp., Madison, WI). For maxipreparations, Qiagen's maxiprep kit (Qiagen Inc., Chatsworth, CA) is used.

Nucleic acids are separated on agarose gels (low EEO, Sigma). For preparative gel electrophoresis, low melt SeaPlaque GTG agarose is used (FMC bioproducts, Rockland, ME), for electrophoretic separation of SSLP genetic markers, Metaphor gels are used (FMC) and

for separation of smaller fragments, as well as for heteroduplex analysis, 10% 19:1 polyacrylamide gels (BioRad) are used. For documentation, nucleic acids separated on gels are photographed on Polaroid black and white print film, iso3000/36° (Cambridge, MA). For blotting, nucleic acids are transferred onto GeneScreen Plus membranes (NEN™ Life Science Products, Boston, MA) or Hybond™-N+ membranes (Amersham, Arlington Heights, IL). Random primer DNA labeling mix is obtained from GibcoBRL Lifescience. Radioisotopes [α-^{32}P] dCTP are delivered by International Biotechnologies Inc. (New Haven, CT). Radioactive signals are visualized on Kodak X-OMAT film (Rochester, NY).

2. Common Media and Buffers:

For the preparation of the buffers and media, double-autoclaved water is used. Prior to use, the solutions are autoclaved or filter-sterilized through 2 µm filters (Nalgene, Rochester, NY).

LB
10 g/l Bacto tryptone
5 g/l Bacto yeast extract
170 mM NaCl

SOC
20 g/l Bacto tryptone
5 g/l Bacto yeast extract
10 mM NaCl
2.5 mM KCl

GM medium (Murashige and Skoog, 1962)
4.39 g/l M&S standard medium
10.0 g/l sucrose
0.5 g/l MES
1 mg/l thiamine
0.5 mg/l pyridoxine
0.5 mg/l nicotinic acid

100 mg/l inositol

YT medium
8 g/l Bacto tryptone
5 g/l Bacto yeast extract
90 mM NaCl
pH to 7.3
15 g/l Agar

YPDL medium
10 g/l Bacto tryptone
5 g/l Bacto yeast extract
170 mM NaCl
20 mM $MgSO_4$
(for plates, supplemented with 15 g agar per liter)

Kings B medium
38 g/l Pseudomonas Agar F
15 ml/l Glycerol
(for plates, supplemented with 15 g agar per liter)

Media are supplemented with antibiotics, if needed: 50 mg/l kanamycin, 50 mg/l ampicillin, 50 mg/l rifampicin, 15 mg/l tetracyclin, 25 mg/l chloramphenicol, or 30 mg/l hygromycin.

IM medium
4.39 g/l M&S standard medium
5.0 g/l sucrose
10 ml/l 6-benzylaminopurine, 1mg/ml solution in 1% KOH

MSE, 10 x
0.2 M MOPS acid
50 mM sodium acetate, pH 5.2
10 mM EDTA
pH to 7.0 with NaOH

TAE buffer, 10 x	48.4 g/l Tris base 11.4 ml/l acetic acid, glacial 10 mM EDTA, pH 8.0
TBE buffer, 10 x	0.89 M Tris-base 0.89 M boric acid 20 mM EDTA, pH 8.0
TE, pH 8.0 (pH 7.5) 1 mM EDTA	10 mM Tris-HCl, pH 8.0 (resp. pH 7.5)
RNA sample buffer	50 ml formamide/bromophenol blue (10:1) 17 ml formaldehyde 10 ml 10 x MSE 0.4 ml ethidium bromide (10 mg/l)
SSC, 20 x	3 M NaCl 0.3 M citric acid, trisodium salt pH 7.0, adjusted with 10 N NaOH
STE buffer	10 mM EDTA 100 mM NaCl 20 mM Tris-HCl, pH 7.5
CTAB buffer	2% w/v CTAB 100 mM Tris-HCl, pH 8.0 20 mM EDTA 1.4 M NaCl 1% w/v PVP 40000

Phenol is saturated with Tris-HCl pH 8.0 except for use in the Trypan Blue stain mix.

3. Biological Materials

A. Plants

Arabidopsis thaliana (Heynh.) ecotypes *Wassilewskija* (*Ws-0*); *Columbia* (*Col-0*), and *Landsberg erecta* (*Ler*) are obtained from Lehle Seeds (Round Rock, TX). A hygromycin resistant NahG line in the *Col-0* background is used for crosses.

B. Microorganisms

E. coli strains

DH5αF' F'/ *end*A1, *rec*A1, *gyrA96*, *thi-1*, *hsd*R17 (rk^-mk^+), *supE44*, *rel*A1, *deo*R, (ϕ80*dlac*Δ(*lacZ*)M15) Δ(*lacIZYA-arg*F) (Hanahan, 1983)

XL1 Blue F'/ $lacI^q$ Δ(*lacZ*)M15, *end*A1, *rec*A1, *gyrA96*, *thi-1*, *hsd*R17 (rk^-mk^+), *supE44*, *rel*A1 (Bullock *et al.*, 1987)

XL2 Blue MRF' F'/ $lacI^q$ Δ(*lacZ*)M15, *end*A1, *rec*A1, *gyrA96*, *thi-1*, *hsd*R17 (rk^-mk^+), *supE44*, *rel*A1, Δ(*mcr*A)183, Δ(*mcr*CB-*hsd*SMR-*mrr*)177 (Stratagene)

DH10B F^- *mcr*A Δ(*mcr*CB-*hsd*SMR-*mrr*) *end*A1, *rec*A1, *gyrA96*, *thi-1*, *hsd*R17 (rk^-mk^+), *supE44*, *rel*A1, *deo*R, (ϕ80*dlac*Δ(*lacZ*)M15) Δ(*lacIZYA-arg*F) (GibcoBRL)

Yeast Strains

Saccharomyces cerevisiae strain AB1380: Matα, *psi+*, *ura3-52*, *trp1*, *ade2-1*, *can1-100*, *lys2-1*, *his5* (Burke *et al.*, 1987)

Agrobacterium tumefaciens

Agrobacterium tumefaciens strain GV3101, containing the pMP90 vir plasmid (Koncz and Schell, 1986)

Pseudomonas syringae strains

Pseudomonas syringae pv. *tomato* DC 3000, Rif^r (Whalen *et al.*, 1991)

Pseudomonas syringae pv. *maculicola* ES4326, Kan^r (Dong *et al.*, 1991)

Peronospora parasitica

pathovar	susceptible *A. thaliana* ecotype	Source
Noco2	*Col-0*	J. Parker, Sainsbury Laboratory, Norwich, England
Emco5	*Col-0*	B. Mauch-Mani, A. Slusarenko, Inst. für Pflanzenbiologie, Zürich, Switzerland
Emwa	*Ws-0*	E. Holub, I. Crute, Horticultural Research Station, E. Malling, England

Erysiphe cichoracearum

Erysiphe cichoracearum strain UCSC is provided by R. Innes (Indiana University, IN).

C. Vectors

Plasmids

pUC19 *colE1 ori*, *lacZ*α, MCS, Amp^r (Yanish-Perron *et al.*, 1985)

pBeloBAC-Kan *ori S*, Kan^r, *rep*E, *par*A *par*B, cosN, MCS, loxP, *lacZ* (Mozo *et al.*, 1998)

pBeloBAC11 *ori S*, Chl^r, *rep*E, *par*A *par*B, cosN, MCS, loxP, *lacZ* (Kim *et al.*, 1996)

pYAC4 *URA3*, *SUP4*, *CEN4*, *ARS1*, *TRP1*, Amp^r, pBR322, *TEL HIS3 TEL* (Burke *et al.*, 1987)

pYAC41, *dito*, but modified restriction sites (Albertsen *et al.*, 1990)

pCIB200 LB Nos*NPTII* RB, *oriT*, *oriV*, Kan^r

Phagemids

pBluescript SKII(+/-) *colE1 ori*, *f1 ori*, *lacZ*, MCS, Amp^r ((Short *et al.*, 1988), and Stratagene)

pHD-1 is identical to pBluescript but contains a polylinker cloned in the *Not*I site (Hofte *et al.*, 1993).

pZL1 *colE1 ori*, *f1 ori*, *lacZ*, *lacI*, *IncA*, loxP, MCS, Amp^r (GibcoBRL)

pCR^{TM}2.1 *colE1 ori*, *f1 ori*, *lacZ*α, MCS, Amp^r, Kan^r (Invitrogen Corp., Carlsbad, CA)

Libraries and stock collections

YAC clones, BAC clones and BAC library filters (TAMU and IGF libraries) are obtained from the ABRC stock center (Ohio State University, OH).

Cosmid library pOCA18 (binary vector, bacterial selection: Tet^r, plant selection: 35S/hyg^r, 10 genome equivalents *Ws-0* inserts, average size 25 kb; Schulz *et al.*, 1994), obtained from ABRC stock center

MSU *est* collection Amp^r, lambda ZipLox (LifeScience BRL) *Sal*I-*Not*I, or lambda ZAPII (Stratagene) *Eco*RI-*Not*I cDNA cloning (0.5 to 6 kb), obtained from the ABRC stock center as excised plasmids.

4. General Laboratory Techniques

All common molecular biological methods, as described for instance in (Ausubel *et al.*, 1987; Sambrook *et al.*, 1989) are performed accordingly and are not described separately.

A. Restriction endonuclease digest of DNA

The endonucleolytic cleavage of DNA by restriction enzymes is carried out according to the manufactures specifications. For CAPS marker development, 5 µl of a 25 µl PCR reaction

is used in a 20 µl digest. For plasmid rescue and Southern blot analysis, including RFLP analysis, digests are performed in 200 µl volume, using up to 3 µg DNA.

B. Purification of DNA fragments

DNA fragments are either gel purified or, for PCR-generated fragments, purified in solution. For gel purification, small pieces of low-melt agarose are isolated and DNA is extracted using the Advantage™ PCR pure Kit (Clontech, Palo Alto, CA). For DNA in solution (e.g. PCR fragments), the GeneClean III kit (Bio101, Inc., Vista, CA) is used according to the manufactures recommendations.

C. Ligation

Ligations of DNA fragments to vector DNA are performed according to standard protocols (Sambrook *et al.*, 1989). Usually, 50 to 100 ng vector DNA are mixed with a two- to threefold excess of fragment DNA in a 30 µl reaction volume. The reactions are performed overnight at 16°C for both sticky and blunt end ligations. One to 5 units T4 ligase are used per reaction. PCR fragments are cloned by TOPO-TA cloning following the instructions of the manufacturer (Invitrogen).

For plasmid rescue, 2 µg of digested genomic DNA are resuspended in 358 µl water, supplemented with 40 µl ligase buffer and ligated in presence of 2 µl T4 ligase (12 Weiss units).

D. Preparation of heat-shock competent *E. coli* cells

A 100 ml LB culture is inoculated with 0.5 ml of a liquid overnight culture of *E. coli* DH5α and grown with shaking at 37°C until an optical density (O. D. $_{600}$) of 0.5 has been reached. Cultures are chilled on ice and cells are collected by centrifugation (5 min, 5000g). Cells are resuspended in 7.5 ml transformation buffer I (100 mM $RbCl_2$, 45 mM $MgCl_2$, 35 mM potassium acetate, 10 mM $CaCl_2$, 0.5 mM LiCl, 15% glycerin, pH 5.8) and incubated for 10 min on ice. Cells are spun down as before and resuspended in 4 ml transformation buffer II (10 mM MOPS, 75 mM $CaCl_2$, 10 mM $RbCl_2$, 15% glyerin, pH 7.0). Aliquots of 100 µl are rapidly frozen and stored at -80°C.

E. Transformation of competent *E. coli* cells

Plasmids are transformed into *E. coli* using a modified version of the heat-shock protocol (Dagert and Ehrlich, 1979). An aliquot of competent cells is thawed on ice and incubated for 10 min with the DNA. After a heat pulse (1 min, 42°C), the cells are again incubated on ice for 2 min. 200 µl SOC media at room temperature are added and the mixture is incubated at 37°C for an hour. Cells are plated on LB plates containing the selective antibiotics and in appropriate cases isopropylthiogalactose (IPTG) and 5-bromo-4-chloro-3-indoyl-β-D-galactose (X-Gal).

For high efficiency transformation, Epicurian Coli® ultracompetent cells (*E. coli* XL-2 Blue; Stratagene) are transformed following the manufactures instructions.

F. Transformation of competent *Agrobacterium* cells

To transform a binary vector into *Agrobacterium*, 40µl electrocompetent *Agrobacterium* cells are thawed on ice and 2 to 10 ng plasmid DNA is added. The mixture is transferred into a prechilled 0.2 ml electroporation cuvette (BioRad) and the cells are electroporated at 2.0 Volts, 600 Ohms, 25 µFarad, 6 msec time constant using a Gene Pulser (BioRad). Immediately, 1 ml of 2 x YT medium is added and the suspension is incubated at 37°C for one hour under shaking. Cells are collected by centrifugation, resuspended in a small volume LB medium and spread onto LB plates containing the appropriate antibiotic (Kanamycin for pCIB200). Plates are inoculated 2 to 3 days at 28°C before inoculating 50 ml liquid LB cultures (supplemented with kanamycin and rifampicin) for transformation. 10 ml of this culture are used after 24 - 36 hours incubation at 28°C to inoculate 500 ml LB cultures. To verify that the plasmid is correctly transformed, plasmids are isolated following a lysis in 5 M NaCl, 20% sarkosyl solution and then the protocol of the Wizard plasmid minipreparation is followed (Promega, section 2.4.7).

G. Plasmid minipreparation

The preparation of small amount plasmid DNA is carried out following a method by Birnboim and Doly (1979). 3 ml overnight cultures are concentrated by centrifugation, and resuspended in 200 µl solution I (50 mM glucose, 25 mM Tris-HCl, pH 8.0, 10 mM EDTA). The solution is placed on ice and 200 µl of solution II (0.2 M NaOH, 1% SDS) are added to

lyse the cells. Cell debris is precipitated with 200 µl 5 M potassium acetate and separated from the supernatant by centrifugation. DNA is precipitated from the supernatant by 1/10 volume sodium acetate and 2 volumes ethanol. The pellet is resuspended in 50 µl water.

Alternatively, the Wizard plasmid minipreparation kit (Promega) is used for 3 ml overnight cultures according to the manufactures recommendations.

H. DNA Sequencing

DNA Sequencing is done according to the Sanger 2',3'-didesoxy technology (Sanger *et al.*, 1977), using the big Dye terminator ready reaction mix (ABI/Advanced Biotechnologies, Inc., Columbia, MD), supplemented with 2 mM $MgCl_2$, 80 mM Tris-HCl, pH 8.0 buffer. For DNA fragments bigger than 5 kb, transposon-mediated sequencing is carried out (Kimmel *et al.*, 1997) using the primer island transposition kit (Perkin Elmer). All sequencing reactions are carried out in Peltier Thermal cyclers (MJ Research Inc., Watertown, MA) and loaded onto 5% acrylamide long-range gels (FMC ready mix). Fluorescence is read by an ABI Prism 377 DNA sequencer (ABI) and bases are called using Phred/Frap/Consed software (University of Washington, Seattle, WA; Ewing *et al.*, 1998; Gordon *et al.*, 1998).

For sequence assembly and comparison, and restriction site mapping, the sequencer software (Gene Codes Corp., Ann Arbor, MI; Version 4.0 for Windows) is used. For similarity searches, BLAST2 software (Altschul *et al.*, 1990) are run at NCBI against GenBank (www/ncbi.nlm.nih.gov/BLAST/) or against the *Arabidopsis thaliana* database (http://genome-www2.stanford.edu/cgi-bin/AtDB/nph-blast2atdb).

5. Plant Care

A. Plant growth conditions

A. thaliana (L.) Heynh. ecotypes *Ler*, *Ws-0* and *Col-0* are sown in 200-ml containers in an all-purpose soil mix (Germination Mix, superfine C. Fafard Inc.; Agawam, MA) that has been autoclaved twice for 70 min or once for 2 hours after 24 h hydration to allow fungal sporulation. The seeds are surface-sterilized with bleach (50% v/v commercial bleach, 0.01% v/v sodium dodecyl sulfate or another wettable agent) for 5 min and for 5 min with 80% ethanol, and washed several times in sterile distilled water before sowing. Plants are grown at

20 - 24°C, 60% relative humidity, 9 hr day/15 hr night (short day, SD), 250 μE/m^2 s. Prior to germination, the flats are covered with plastic domes. For older plants, the soil surface is allowed to dry between waterings. Alternatively, plants are cultivated on GM-agarose in petri dishes in 0.1% GM top agar under sterile conditions and either SD or long day (LD; 15 hr day, 9 hr night) in high densities (up to 10 plants per 1 cm^2).

B. Pollination

Cross-pollination of *A. thaliana* is performed on half-closed buds of flowers from the female parent plant. It is confirmed with the aid of a dissecting microscope that the anthers have not yet released pollen on the stigma. From the male parent plant, a dehiscing anther is removed with forceps and pollen is transferred to the stigma of the female parent.

C. Transformation of *Arabidopsis thaliana* by *Agrobacterium*

Plants are genetically transformed using an adapted protocol of vacuum infiltration (described by Bechtold *et al.* (1993)). *A. thaliana* plants, which have bolted and just started flower formation (4 - 5 weeks old) are immersed in infiltration media (IM) containing a washed 3 liter late log phase culture of *A. tumefaciens* GV3101 (O.D. $_{600}$ = 0.7 or above), previously electrotransformed with a binary vector (section 4F). Infiltration is accomplished by creating and releasing a vacuum in the chamber containing the plants. Plants are then cultured as described above and the seeds are harvested and subjected to selection either on GM plates containing 50 mg/l Kanamycin (Valvekens *et al.*, 1988) or on soil with 3 to 4 spray treatments of 160 mg/l Basta® (glufosonate ammonium) in the first two weeks of development in order to identify the transformants (Akama *et al.*, 1995).

D. Treatment of plants with chemical inducers of SAR

Plants growing on soil are sprayed with SA (5 mM), INA (0.375 mM) or BTH solutions (0.1 – 5.0 mM, usually 0.3 mM) containing 0.01% v/v of a surface-active agent (SilWet L77). Plants cultivated on agar plates are exposed permanently to the inducers mixed in the GM-medium at different concentration ranges (BTH up to 250 μM, INA up to 300 μM and SA up to 150 μM final concentration in the plate).

E. Luciferase activity determination *in vivo*

To monitor luciferase activity on living plants, light emission is measured in a photon imaging device (Hamamatsu Inc. Tokyo) equipped with an ARGUS 50 photon-counting image-processor 15 min after the plants are sprayed with a luciferin solution (7.5 mM luciferin, 0.1% v/v Silwet 77). The most sensitive regulation and intensified CCD microscope video cameras (C2400-25 VIM) are used during signal integration for 10 min at 20°C. Plants that gave a positive measurement of light emission are identified, isolated and cultivated for further investigations.

6. Pathogen Treatment

A. Treatment with and culture of *Peronospora*

To maintain fungal stocks, weekly transfers of *P. parasitica* are made. Transfers are accomplished by spraying oospores on compatible *A. thaliana* cultivars that are grown under high relative humidity (95%) at 15°C in a culture chamber.

For pathogenicity tests, *Peronospora* oospores are isolated from infected *A. thaliana* leaves by vortexing the leaves in distilled water. Spores are counted in a hemacytometer and the concentration is adjusted to 10^5 - 10^6 spores per ml. The supernatant is then used directly either to spray the planosphaere (Dietrich *et al.*, 1994) or to infiltrate the leaf apoplast of *Arabidopsis thaliana* with a 1 ml syringe gently pressed onto the subfacial leafside.

B. *Erysiphe* inoculations

Resistance to *Erysiphe cichoracearum* strain UCSC is tested by brushing sporulating *Col-0* plants onto four weeks old plants, as described by Frye and Innes (1998). Disease is scored 10 days after inoculation on at least 10 plants per mutant, according to the following rating: rating 1: 0-1 leaf per plant showed hyphal growth; rating 2: 2 - 4 leaves per plant showed hyphal growth; rating 3: more than 4 leaves per plant showed hyphal growth. Mean values and standard deviations are calculated for one experiment. The experiments are repeated three times.

C. *Pseudomonas* infections

For the analysis of resistance to compatible phytopathogenic bacteria, the apoplast of leaves of four weeks old *cim* plants, BTH-activated *Col-0* (0.3 mM, 2 days prior to infection) and water-treated *Col-0* control plants are injected with *Pseudomonas syringae* pv. *maculicola* ES 4326 (Schott *et al.*, 1990) or *Pseudomonas syringae* pv. *tomato* DC3000 (Dong *et al.*, 1991) at 2 x 10^5 cells per ml. Samples are taken at 0, 1, 3 and 5 days after injection. For each time point, four leaf punches are pooled, ground in 10 mM $MgCl_2$ and plated in appropriate dilutions on Kings B medium supplemented with rifampicin (50µg/ml), or kanamycin (100 µg/ml). Standard deviations are calculated from four independent experiments. The significance of differences between mean values is evaluated by Student's *t* test. Differences are considered to be significant at $P > 0.6$.

To monitor the response of *cim* mutants to avirulent *Pseudomonas* strains, high concentrations (5 x 10^7 cells per ml) of *Pseudomonas syringae* pv. *tomato* DC3000 *avrPt2* or *avrRpm1* (cloned into pVSP61) are infiltrated into leaves and symptom development is compared to wild-type and BTH pretreated plants over a period of 48 hours.

Before each pathogenicity test, *cim* mutants are selected based on *in vivo* expression of the *PR-1*/luciferase gene. For crosses of *cim* mutants to the NahG line, where *PR-1*/luciferase expression is suppressed in all cases, resistance is evaluated on population level.

7. Staining Methods

A. Trypan Blue staining for cell death and fungal infection

In order to detect lesions and fungal hyphae on leaves, a protocol by Keogh (Keogh *et al.*, 1980) is adapted for *Peronospora* (Koch and Slusarenko, 1990). Leaves are covered with prewarmed (65°C) Trypan Blue-stain mix (9.3 ml phenol, 10 ml lactic acid, 10 ml glycerol, 10 ml Trypan Blue (1 mg/ml)). The dye is infiltrated by repeated vacuum application and release followed by brief boiling in a microwave. After an one-hour incubation at room temperature, leaves are cleared in a nearly-saturated (250% w/w) chloral hydrate solution overnight and slides are prepared the next day.

B. 3,3'-dihexylocarbocyanine iodide [$DiOC_6(3)$] staining for *Erysiphe*

To visualize *Erysiphe* infection and fungal structures, a fluorescence dye staining is performed on infected leaves (Duckett and Read, 1991). Leaves are incubated for 2 min in 50 µg/ml [$DiOC_6(3)$] stain (Sigma Chemical Corp.), cleared for 30 sec in distilled water and mounted in water under a coverslip. Fluorescent fungal hyphae are detected at 520 nm after blue light excitation (450 - 490 nm) with an epifluorescence microscope (Leitz, Wetzlar, Germany).

C. Callose staining

Callose depositions are detected using an aniline blue stain on 5 µm thick leaf sections (Hunt *et al.*, 1997). Leaves are fixed in 10% formaldehyde solution (45% ethanol, 10% formaldehyde, 5% acetic acid), and embedded in paraffin blocks. Microtom leaf sections (made by Experimental Pathology Laboratory, Durham, NC) are mounted on microscope slides and deparaffinated by two successive 5 min incubations in 100% xylene, two 5 min incubations in 100% ethanol and one 5 min incubation in each of the following: 75% ethanol, 40% ethanol, and water. For callose staining, samples are incubated for 5 min in 0.15 M K_2HPO_4 and 0.01% aniline blue. Samples are mount in 70% glycerol, 30% aniline blue stain and visualized using ultraviolet epifluorescence (390 - 430 nm) as described in Dietrich *et al.* (Dietrich *et al.*, 1994).

8. Biochemical Methods

A. Luciferase *in vitro* assay

A modification of the extraction method described by Millar *et al.* (1992) is used to assay luciferase activity *in vitro*. Plant tissue is ground in liquid nitrogen, resuspended in 500 µl prechilled grinding buffer (10 mM phosphate buffer, pH 7.5, 0.1 mM DTE), and centrifuged for 10 min (1000g, 4°C). 100 µl of the supernatant are mixed with 100 µl luciferin substrate (Promega Corp., Madison, WI or Biosynth International Inc., Naperville, IL) and light emission is integrated for 10 sec by a luminometer (Monolight 2010, Analytical Luminescence Laboratory; Ann Arbor, MI).

To determine the total protein content in the samples, the absorbence of 5 µl of the supernatant is measured at 595 nm after the addition of 20% v/v protein assay-solution (BioRad; Hercules, CA) in a total volume of 500 µl.

B. PAL enzyme activity assay

0.5 g tissue, in presence of 50 ?g Dowex 1X2-200 (Sigma Chemical Corp.) and quartz sand are ground in liquid nitrogen. After addition of 1 ml extraction buffer (0.2 M Tris-HCl, pH 7.8, 14 mM β-mercaptoethanol), the solution is incubated 10 min on ice and the cell debris is separated by centrifugation at 4°C (10 min, 20,000g). For the enzyme assay, 100 µl of the supernatant are mixed with 400 µl 0.1 M Tris-HCl, pH 8.8, and 500 µl 20 mM phenylalanine (Sigma Chemical Corp.) dissolved in Tris-HCl, pH 8.8. PAL deaminates phenylalanine to cinnamic acid whose concentration can be measured by absorption spectroscopy at 290 nm (λ_{max} at 275 nm; $\varepsilon = 20000\ cm^{-1}M^{-1}$). The reaction is incubated at 30°C, and the O. D. $_{290}$ is measured at 10 min, 30 min, 60 min and 90 min after reaction start. PAL activity in µKat/kg protein can be calculated according to: ΔE/h x 27.8/mg protein (Kombrink and Hahlbrock, 1986). Protein content is determined as described above (8A) using the Bradford reagent.

C. Salicylic acid extraction

In order to measure the content of free salicylic acid and its monosaccharide conjugates, samples are harvested in triplicates and analyzed as previously described (Enyedi *et al.*, 1992; Uknes *et al.*, 1993). 0.3 g of frozen, ground *A. thaliana* leaves are extracted with 3 ml 90% methanol during 20 minutes of sonication. After centrifugation at 4000g for 20 min, the pellet is further extracted with 2 ml 100% methanol and then spun down again (4000g, 20 min), and the two supernatants are combined. Samples are split into two equal parts and dried in a speedvac. The first series of samples (free SA) is suspended in 2.5 ml 5% trichloroacetic acid (TCA; 5 min sonication) and SA is extracted twice with 2.5 ml extraction buffer. The second series (total SA) is first subjected to hydrolysis with 40 units β-glucosidase (Sigma Chemical Corp.) in 0.1 M sodium acetate (pH 5.2) at 37°C for 1.5 hours. 2 ml 5% TCA is then added and the acidic phase is extracted three times with 2.5 ml extraction buffer (ethylacetate : cyclopentane : isopropanol = 100 : 99 : 1). Each extract is dried in a speedvac, resuspended in 150 µl 20% methanol, filtered in a spincolumn (Titan-MSF nylon microsample filters, 0.2

µm; SRI Scientific resources Inc.; Eatontown, NJ) and transferred to an HPLC autosampler vial. For chromatographic separation and measurements, 50 µl of each extract is injected in a C-18 HPLC column (Dynamax 60, Rainin Instrument Comp.; Woburn, MA). SA is eluted at 44°C in a 15 min linear gradient (solution A: 2 mM aqueous sodium acetate plus 5% - 30% solution B: 2 mM sodium acetate in 70% methanol, followed by 100% solution B and reequilibration for 5 min in 5% solution B) and detected by a fluorescence detector at the column end after excitation at $\lambda = 295$ nm. Area values are converted to concentrations using a standard curve.

D. Camalexin extraction

For camalexin extraction, triplicates of 3 leaf discs each (Cork borer #1) are pooled and extracted in hot methanol according to Dr. J. Glazebrook (http.//genome-www.stanford.edu/*Arabidopsis*/cshl-course/8-defense.html). Extracts are run on a Si250 TLC plate (J. T. Baker) in ethyl acetate : hexane (87:13) mobile phase. Tissue induced by *Pseudomonas syringae* pv. *maculicola* ES 4326 harvested 5 days after inoculation served as a control for extractions, synthetic camalexin (provided by Dr. T. Maetzke, Novartis Inc.) served as a standard during thin layer chromatography. Camalexin is detected under a long-wave ultraviolet lamp.

9. Extraction Of Nucleic Acids

A. Extraction of plant RNA

Total RNA is isolated from 1 g frozen, powderized leaf tissue that is ground to a fine powder in liquid nitrogen. The samples are resuspended in 2.5 ml RNA extraction medium (50 mM Tris-HCl, pH 8.0, 4% w/v *p*-amino salicylic acid, 1% w/v 1,5-naphtalene disulfonic acid (Arcos Chemicals, NJ) and 2.5 ml water-saturated phenol (Lagrimini *et al.*, 1987). After addition of 2.5 ml chloroform, phases are separated by centrifugation (10 min at 7000g). The aqueous phase is transferred to a new tube and nucleic acids are precipitated with the addition of 1/10 volume 3 M sodium acetate, pH 5.2 and 2 volumes ethanol at -20°C for 30 min. Precipitates are spun down (10 min at 7000g) and the dried pellets are resuspended in 2 ml double-distilled water. RNA is precipitated overnight at 4°C with the addition of 1.25 ml 8 M

LiCl. The precipitate is pelleted by centrifugation (10 min at 7000g) and the pellet is rinsed with 80% ethanol. RNA pellets are resuspended in 100 µl water and the absorbence at 260 nm and 280 nm are measured in a spectrophotometer (UV-160 visible recording spectrophotometer, Shimadzu; Columbia, SC) to determine the amount and the purity of the RNA.

For one-leaf extractions, a modification of the above protocol is used. The leaf is ground in liquid nitrogen, and 500 µl of an 1:1 mixture phenol : extraction buffer (0.1 M LiCl, 100 mM Tris-HCl, pH 8.0, 10 mM EDTA, pH 8.0, 1% SDS) at 80°C are added. After adding 250 µl chloroform, the phases are separated by centrifugation and nucleic acids in the aqueous phase are precipitated with sodium acetate and ethanol at -20°C. The pellet is resuspended in 8 µl water.

B. PolyA RNA purification

PolyA RNA is enriched using a poly dT magnetic bead technique (Promega). One mg total RNA is hybridized to the poly dT-biotin nucleotides according to the manufactures recommendations. Avidin-coated magnetic beads allowed the separation of polyA RNA from non poly-adenylated RNA. The RNA is released from the poly dT probes in low salt buffers and concentrated by precipitation with ethanol. 10% of the obtained polyA RNA is used for spectrometric analysis.

C. Plant DNA minipreparation

Plant DNA is extracted using the CTAB-method described by Rogers and Bendich (1988). 1 - 2 leaves per sample are ground in liquid nitrogen with a Polytron (Brinkmann Instruments Inc. Westbury, NY), then vortexed with 200 µl 2 x CTAB buffer. After heating at 65°C for 15 min, 200 µl chloroform are added and the well-mixed extraction is centrifuged for 2 min (10,000g). DNA is precipitated from the resulting supernatant with 3 volumes of ethanol at -20°C. The precipitate is spun down (10,000g, 15 min) and the pellet is rinsed with 70% ethanol. After drying, the pellet is resuspended in 30 µl 10 mM Tris-HCl, pH 8.5.

D. Plant DNA extraction for Southern blot analysis and pooled progeny analysis (F3 populations)

1 to 2 g ground tissue are mixed on ice with 12 ml extraction buffer (0.1 M Tris-HCl, pH 8.0, 50 mM EDTA, pH 8.0, 0.5 mM NaCl, 10 mM β-mercaptoethanol; Dellaporta *et al.*, 1983). After adding 0.8 ml 20% SDS, the extract is incubated at 65°C for 10 min. Cell debris is precipitated with 4 ml 5 M potassium acetate during a 20 min incubation at 4°C and separated from the supernatant by centrifugation at 4°C (10 min, 8000g). The supernatant is filtered through prewetted Miracloth and DNA is precipitated with 8 ml isopropanol at -80°C (30 min). DNA is pelleted, dried and resuspended in 4 ml TE and again precipitated with sodium acetate and ethanol. After resuspension in 400 µl TE, an RNase A digest is performed for 10 min at 37°C (final concentration RNase A: 50 µg/ml). The samples are extracted once with a 1:1 mixture of phenol : chloroform, once with chloroform, precipitated by addition of sodium acetate and ethanol and resuspended in 50 µl water. For less tissue (15 - 500 mg), the protocol is scaled down 20-fold and the RNase treatment is omitted. 10 µl is used in restriction digests for Southern blot analysis.

E. BAC DNA minipreparation

BAC DNA minipreparations are done according to a protocol by Sinnett *et al.* (1998), using a modified alkaline lysis method. 3 ml LB overnight cultures containing either 50 ?g/ml kanamycin (IGF BACs) or 12.5 µg/ml chloramphenicol (TAMU BACs) are pelleted and resuspended in 100 µl of chilled resuspension solution (25 mM Tris-HCl, pH 8.0, 50 mM glucose, 10 mM EDTA, pH 8.0). Cells are lysed by adding 200 µl lysis buffer (0.2 *N* NaOH, 1% SDS). Cell debris and chromosomal DNA are precipitated by 150 µl 5 M potassium acetate, pH 4.8. After a 5 min centrifugation at maximal speed in a tabletop centrifuge, the supernatant is transferred to a new tube and the crude DNA is precipitated by adding 2 volumes of ethanol. The DNA is pelleted for 5 min as before, washed in 70% ethanol and resuspended in 100 µl of TE buffer containing 0.1% SDS and 100 µg/ml proteinase K, followed by a one-hour incubation at 37°C. The reaction is extracted with 100 µl phenol : chloroform (1:1), then with 100 µl chloroform. The DNA is precipitated with 2 volumes ethanol and washed as before, then resuspended in 50 µl water.

F. BAC DNA maxipreparations

To obtain bigger quantities of BAC DNA, a protocol provided by Choi *et al.* (1995) is used in a modified version. 2 liter overnight bacterial cultures (LB plus antibiotic) are harvested by centrifugation. The pellet is resuspended in a lysozyme solution (50 mM glucose, 10 mM EDTA, 25 mM Tris-HCl, pH 8.0, 5 µg/ml lysozyme) and incubated on ice for 5 min. 40 ml of an alkaline lysis solution (0.2 *N* NaOH, 1% SDS) are added and after 5 min incubation at 4°C, 30 ml of ice-cold potassium acetate solution (5 M, pH 4.8) are added. After freezing at -80°C, the precipitate is pelleted (15 min, 15000g), the supernatant is filtered through miracloth, and DNA is precipitated by 0.6 volumes isopropanol at -80°C. DNA is pelleted (30 min, 15000g), dissolved in TE and subjected to RNase A digest (20 µg/ml, 45 min at 37°C). DNA is extracted with an equal volume phenol : chloroform (1:1), and with chloroform, then precipitated with sodium acetate and ethanol. DNA is taken up in 200 µl water.

Alternatively, the purification followed the Qiagen maxipreparation protocol for very-low-copy cosmids (Qiagen Inc., Valencia, CA). The DNA is taken up in 20 µl of 10 mM Tris-HCl, pH 8.5.

G. YAC DNA preparation

To isolate yeast DNA, a protocol by Hoffman and Winston (1987) is used. Cells of a stationary 5 ml YEPD liquid culture are collected by centrifugation, washed in water and lysed in lysis buffer (2% Triton X-100, 1% SDS, 100 mM NaCl, 10 mM Tris-HCl, pH 8.0, 1 mM EDTA, pH 8.0) in presence of 0.2 ml phenol : chloroform (1:1) and 0.2 ml glass beads under intensive vortexing. After adding 0.2 ml TE and centrifugation, the aqueous phase is extracted with chloroform and DNA is precipitated with two volumes ethanol. The pellet is resuspended in 0.4 ml TE and the DNA is precipitated with 10 µl ammonium acetate and 1 ml ethanol. DNA is resuspended in 50 µl TE.

10. Analysis Of Membrane-Bound Macromolecules

A. Analysis of RNA

For RNA blot analysis, 10 µg of purified RNA (0.25 µg poly A RNA) in 15 µl water is heated for 15 min at 65°C after addition of 34 µl RNA sample buffer (65% v/v formamide:bromophenol blue (1:10), 21.5% v/v formaldehyde (37%), 13% MSE, 0.5% v/v ethidium bromide (10 mg/l)). The samples are loaded on a denaturing agarose-gel (1.2% w/v agarose, 1 x MSE, 3% v/v formaldehyde) and run at 12 V/cm for 1 hour (Sambrook *et al.*, 1989). The gel is blotted overnight in 6 x SSC (Ausubel *et al.*, 1987) on a GeneScreen Plus® membrane, and the efficiency of the RNA transfer is checked by viewing the blot under UV illumination. RNA is linked to the membrane by UV-crosslinking (Stratalinker®, Stratagene; La Jolla, CA) at 1200 µJ. After at least 1 hour prehybridization at 65°C (hybridization-buffer: 500 mM $NaPO_4$, pH 7.0, 1 mM EDTA, 7% v/v SDS, 1% w/v BSA (fraction V); Church and Gilbert, 1984), 100 µl of an [α-^{32}P] dCTP labeled probe (random priming, Feinberg and Vogelstein, 1983) is added and the membrane is incubated overnight; Church and Gilbert, 1984). The membrane is washed twice with washing buffer (40 mM $NaPO_4$ buffer, pH 7.0, 1mM EDTA, 1% v/v SDS) containing 5 g/l bovine serum albumin (BSA) (each 20 min, 65°C) and once without BSA (15 min, 65°C). The blot is exposed for at least half an hour to a PhosphorImager® screen (Molecular Dynamics; Sunnyvale, CA) and, depending upon the observed intensity, exposed for several hours to days on XAR-5 scientific imaging film at -80°C in the presence of an intensifier screen. The experiments are repeated at least twice for every probe.

Table 4: Clones Used As Probes For The Characterization Of Gene Expression In *cim* Mutants In Northern Blot Analysis And DNA Microarrays.

Probe	Clone Name	Obtained From	Reference
PR-1, PR-2, PR-5, PR-4	pathogenesis related protein cDNAs	8	(Potter *et al.*, 1993; Uknes *et al.*, 1992)
PDF1.2	plant defensin *est* 37F10T7	1	(Penninckx *et al.*, 1996)
THI2.1, 2.2	thionin cDNAs	1	(Epple *et al.*, 1997)
LTP1	lipid-transfer protein1 cDNA	2	(Thoma *et al.*, 1993)
LOX1	lipoxygenase cDNA CD4-16	3	(Melan *et al.*, 1993)
PAL1	phenylalanine ammonia lyase 1 gene	3	(Wanner *et al.*, 1995)
CHS	chalcone synthase, *est* ATTS0223	3	(Shirley *et al.*, 1995)
MLO	*mlo* homolog *est* 205N12T7	3	(Buschges *et al.*, 1997)
NIM1	nim (no inducible immunity) cDNA	8	(Ryals *et al.*, 1997)
NDR1	non-race-specific disease resistance *est* 96E6T7	3	(Century *et al.*, 1997)
DAD	defense against death cDNA 127H23T7	4	(Sugimoto *et al.*, 1995)
MOM	two-hybrid interactor of *DAD*	4	J. Dangl, unpublished
LSD1	lesion simulating disease resistance cDNA	4	(Dietrich *et al.*, 1997)
LLS1	lethal leaf spot homolog, *est* 84E8T7	3	(Gray *et al.*, 1997)
MNSOD, CUZNSOD, FESOD	superoxid dismutases, *est* 105G4T7, *est* 92L6T7, *est* 34D9T7	3	(Bowler *et al.*, 1992)
CAT2, CAT3	catalase 2 and 3, *est* 38C1T7, *est* 40A1T7	3	(Frugoli *et al.*, 1996)
PERXC	peroxidase C, *est*B25XP	4	(Jabs *et al.*, 1996)
CPER	cationic peroxidase, *est* 40H1T7	4	(Jabs *et al.*, 1996)
GST	glutathione-S-transferase typeIII, *est* 90B13T7	2	(Marrs, 1996)
β-TUBULIN	PCR fragment 689 to 1136 bp	8	(Delaney *et al.*, 1994)
RAB18	responsive to abscisic acid cDNA	5	(Gosti *et al.*, 1995)
PAT1	phosphoribosyl-anthranilate synthase gene	6	(Rose *et al.*, 1992)
GluPER	glutathione peroxidase, *est* 139F9T7	3	(Eshdat *et al.*, 1997)
SPS	sucrose-phosphate synthase cDNA	8	D. Guyer, unpublished
PROTOX	protoporphyrin oxidase, cDNA	8	(Lermontova *et al.*, 1997)
EPSPS	5-enolpyruvylshikimate 3-phosphate synthase cDNA	8	(Klee *et al.*, 1987)
APX	Ascorbate peroxidase, cDNA	7	(Kubo *et al.*, 1992)
GS	Glutamine synthetase, cDNA	7	(Peterman *et al.*, 1991)
ATVSP	vascular storage protein, *est* ATTS0751	3	(Berger *et al.*, 1995)
CXC750	pathogen-inducible *est* 31G11T7	3	(Aufsatz and Grimm, 1994)
BIO2	biotin synthase, cDNA	7	(Patton *et al.*, 1996)

1 P. Epple and H. Bohlmann, ETH Zürich, Switzerland
2 A. Molina, Escuela Tecnica Superior de Ingenieros Agronomos, Madrid, Spain
3 ABRC stock center, Columbus, OH
4 J. Dangl, University of North Carolina, Chapel Hill, NC
5 J. Giraudat, Institut des Sciences Végétales, CNRS, Gif-sur-Yvette, France
6 A. Rose, University of California, Davis, CA
7 D. Patton, Novartis Crop Protection, Inc., NC
8 Genes cloned in J. Ryals' laboratory

B. Southern blot analysis

Approximately 0.5 to 5 µg DNA is digested overnight in 200 µl reaction volume. DNA is precipitated by sodium acetate and ethanol and resuspended in 20 µl water. To separate the restriction fragments, the DNA is loaded in presence of 4 µl loading dye onto a 0.9% TBE agarose gel (run either overnight at 1 V/cm, or 3 to 4 hours at 4 V/cm). The gel is soaked in 0.25 *N* HCl for 20 min, then for 30 min in denaturation solution (1.5 M NaCl, 0.5 M NaOH) and twice for 20 min each in neutralization solution (3M NaCl, 0.5 M Tris-HCl, pH 7.5), as described by Ausubel *et al.* (1987). The DNA is transferred overnight onto a nitrocellulose GeneScreen Plus® membrane (Du Pont-New England Nuclear) in 10 x SSC, as described above for Northern blots (section 10A). Hybridization and washing are performed as for Northern blot analysis.

C. Dot blot analysis

Dot blots are performed using a Bio-Dot apparatus (Bio-Rad) as described by Görlach *et al.* (1995). Two µg of total RNA are denatured in 6 x SSPE (20 x SSPE: 20 mM EDTA, pH 7.4, 3 M NaCl, 0.2 M sodium phosphate, pH 7.4) containing 20% deionized formaldehyde for 15 min at 55°C and then chilled on ice. Two volumes of ice-cold 15 x SSPE are added and the samples are applied to a GeneScreen Plus® membrane which is pretreated with 12 x SSPE. After crosslinking, prehybridization and hybridization are performed as described for Northern blot analysis (section 10A).

D. Hybridization of BAC filters and colony filters

BAC filters are hybridized according to the TAMU BAC filter manual (Version 2, http://tamu.edu:8000/~creel/bacman2.html). New filters are prehybridized at 65°C twice for 8 hours in prehybridization buffer (0.5 M $NaHPO_4$, pH 7.2, 7% SDS, 1% BSA, (fraction V), 1

mM EDTA, 10 mg/ml sheared salmon sperm DNA), used filters only once. After adding the probe, hybridization took place in the same buffer for 18 to 36 hours at 65°C. Filters are washed with 0.5 x SSC, 0.1% SDS 3 times for 20 min at 65°C and exposed on a PhosphorImager screen (MolecularDynamics) as described above.

E. Western blot analysis

For Western (protein) blot analysis (Towbin *et al.*, 1979), proteins from 0.25 g pulverized tissue are extracted in 500 µl extraction buffer (0.25 mM Tris-HCl, pH 6.8, 4.5 M Urea, 2% SDS, 5% µ-mercaptoethanol). Equal amount of protein (determined by Bradford reagent, section 2.8.1) are loaded onto a 10% Tris-glycine gel (Novex, San Diego, CA) and run at 10 V/cm for 1 to 2 hours in 1 x running buffer (25 mM Tris-HCl, pH 8.3, 250 mM glycine, 0.1% SDS). The gel is soaked for one hour in equilibration buffer (20 mM Tris-HCl, pH 8.0, 150 mM glycine, 20% methanol) and proteins are transferred by electroblotting onto a nitrocellulose membrane (100 V, 1 hour constant current, Novex X-Cell II Blot module). The membrane is washed for 10 minutes in wash buffer (1 x PBS, 0.1% Tween-20), then incubated for one hour in blocking buffer (1 x PBS, 0.1% Tween-20, 5% milk powder). The PR-1 specific antibody is bound at 4°C overnight in incubation buffer (1 x PBS, 0.1% Tween-20, 1% milk powder) and unspecifically bound antibodies are removed by washing the membrane four times in wash buffer. The second antibody (antiIgG rabbit conjugated with alkaline phosphatase) is bound for 3 hours at room temperature in incubation buffer and the membrane is washed as before. The protein is detected using the NBT (*p*-nitroblue tetrazolium) method as described by Harlow and Lane (1988).

11. PCR Protocols And Genetic Markers

A. SSLP PCR conditions

Polymerase chain reactions (PCR) for microsatellite amplification (SSLP) are carried out following a protocol of Bell and Ecker (1994) in a GeneAmp PCR System 9600 or GeneAmp PCR system 9700 (Perkin Elmer-Roche Corp. Foster City, CA) using a final volume of 25 µl. Each reaction contained 5 µl of 10 fold diluted DNA from the DNA minipreparation (see above, section 2.9.3), 2.5 µl PCR 10 x reaction buffer (Perkin Elmer), 2

µl of a 10 mM dNTP stock solution, 1 µl of forward and reverse primers (20 ?M primer stocks), 0.3 µl AmpliTaq® gold DNA polymerase (5 U/ml, Perkin Elmer) and water *qs.* 25 µl. A typical reaction temperature cycle is: 10 min at 94°C, 40 cycles of 15 sec 94°C, 15 sec 55°C and 30 sec 72°C, then a last polymerization step at 72°C for 10 min. 3 µl loading dye are added and 15 µl of each reaction mix are loaded on a 4% metaphor intermediate melting agarose TBE gel (containing 2.5 x 10^{-3}% v/v ethidium bromide). Gels are run for 2 h at 125 V in 1x TBE. Bands are observed under UV-illumination and documented with Polaroid 3000/36° instant film.

B. CAPS marker PCR conditions and other PCRs

For CAPS markers, the PCR conditions are more variable depending on the length of the expected fragment size (2 min for 2 kb) and the melting temperature of the primers. A typical thermocycle program is: 10 min at 94°C, 35 cycles of 30 sec 94°C, 30 sec 56°C and 2 min 72°C, then a last polymerization step at 72°C for 10 min (Konieczny and Ausubel, 1993). If PCR beads are used (Pharmacia Biotech), only 2 min at 94°C are used prior to thermo cycling. Pooled PCR samples are used for restriction digest of fragments. For sequenced PCR fragments, restriction fragment polymorphisms are identified using the dCAPS software (Neff *et al.*, 1998) rather than by random trial.

For sequencing, PCR fragments are pooled from at least two PCR reactions to minimize sequence differences generated during PCR, purified in solution (section 2.4.2) and diluted to a concentration of 100 µg/ml. Alternatively, PCR fragments are TA-cloned into pCR2.1-TOPO following the instructions given by the manufacturer (Invitrogen).

C. Linkage analysis

A set of 30 SSLP and CAPS primer pairs (see Appendix) is used on segregating F2 populations to establish an initial map position. Genetic map distances are determined using MAPMAKER 3.0 b (Lander *et al.*, 1987; Lincoln *et al.*, 1992) run on a Sun SPARC workstation. Recombination frequencies are calculated using the MAPMAKER F2 algorithm and converted to map distances in centiMorgans (cM) using the Kosambi function (Kosambi, 1944).

D. Long range PCR

Long range PCR is performed on 0.5 µg genomic DNA (Mundy *et al.*, 1995). The lower phase (40 µl) contained final concentrations of 1 x buffer (Perkin-Elmer), 200µM of each dNTP, 1.25 mM MgOAc, and 1.5 µM of each primer. It is covered with a wax bead and heated at 80°C for 5 min, then chilled to 20°C. The upper phase (60µl), containing 450 pM DNA, 4 Units *rTth* polymerase (Perkin Elmer) and 1 x buffer, is added on top of the lower phase. Thermocycling is as follows: 94°C, 1 min, 16 cycles (94°C 30 sec, 68°C 10 min), and 14 cycles (94°C 30 sec, 68°C 10 min with 15 sec extension every cycle), 72°C, 10 min. 10 µl of the reactions are analyzed on 0.9% agarose gels.

E. Other techniques for polymorphism detection

To identify single nucleotide polymorphisms in short DNA fragments (up to 1 kb), two additional techniques are used, the heteroduplex analysis on high resolution gels (Hauser *et al.*, 1998) and the RNase cleavage assay (Myers *et al.*, 1985).

For heteroduplex analysis, equal amounts (20 - 50 ng) of PCR fragments from two different *A. thaliana* ecotypes are denatured at 95°C for 5 min and cooled slowly to 42°C in the heating block. 5 µl of the samples are mixed with 1 µl triple dye (xylene cyanol, bromophenol blue, orange G; FMC Bioproducts), and applied onto a 0.7 mm thick 1x MDE gel (FMC Bioproducts), containing 2.5 M urea. Electrophoresis is performed in 0.6 x TBE buffer using an adjustable slab gel unit (CBS Scientific, Del Mar, CA) with constant voltage (15 V/cm). After the run, the homo- and heteroduplexes are visualized by staining the gel with ethidium bromide. As a positive control, marker F21i5sp6, as described by Hauser *et al.* (1998) is used.

For RNase mismatch detection, genomic fragments (200 bp - 1 kb) are amplified from two *A. thaliana* ecotypes using PCR primers with incorporated T7 (5'taatacgactcactataggg - SEQ ID NO:1) and SP6 (5'atttaggtgacactatagga - SEQ ID NO:2) promoters. Using the Sp6 and T7 RNA polymerase, both sense and antisense RNA probes are made, according to the manufacturers instructions of the MisMatch Detect™ II kit (Ambion, Inc., Austin, TX). Equal volumes of SP6 transcripts are mixed to T7 transcripts of the other ecotype, heated at 95°C for three min and cooled to room temperature. Different RNase digestions of the homo- and

hetero-RNA duplices are performed as recommended. Digestion products are analyzed on a 2% agarose gel (5 V/cm, 1 x TBE, 25 min).

F. Preparation of BAC ends by inverse PCR

To generate probes from BAC ends, we modified a protocol described by Mozo *et al.* (1998). 1/10 of a pBelo BAC DNA minipreparation is digested with either *Eco*RV, *Hin*cII, *Rsa*I or *Eag*I (T7 end), or with *Hha*I, or *Hae*III (Sp6 end). The digest is extracted with phenol : chloroform (1:1), and precipitated with sodium acetate and ethanol. Ligation is performed in a 100 µl volume for at least three hours at room temperature. Ligase is heat inactivated at 70°C (15 min), DNA is precipitated by ethanol, and circularized DNA is cleaved with *Pvu*I (T7 end) or *Bsr*BI (Sp6 end) in a 10 µl volume reaction. For amplification of the pBelo flanking DNA, standard PCR is performed with 56°C annealing temperature and 2 min extension time at 72°C, using for the T7 end the primers: 5'ttcccaacagttgcgcagc (SEQ ID NO:3) and 5'tcttcgctattacgccagct (SEQ ID NO:4), and for the Sp6 flanking DNA, the primers: 5'tcacacaggaaacagctat (SEQ ID NO:5) and 5'acacaacatacgagccggaa (SEQ ID NO:6). PCR fragments are purified as described in section 2.4.2 and sequenced.

G. Thermal asymmetric interlaced PCR of T-DNA insert junctions

To clone the flanking genomic DNA of T-DNA inserts, thermal asymmetric interlaced PCR (TAIL PCR) is used as described by Liu *et al.* (1995). One out of six low stringency primers is used in successive PCR with three nested high stringency primers on either the right border, or the left border of the T-DNA. Reactions are performed on 5-fold diluted CTAB DNA minipreparations. The products of the second and third PCR are analyzed on agarose gels. If a small size difference between the second and the third PCR fragment is detected, the product of the third PCR is either purified for direct sequencing or cloned for sequencing by TA cloning into the vector pCR2.1.

12. Genomics technologies

A. Arraying libraries

Est stock cultures are duplicated from liquid cultures by transferring cells into 96 well flat bottom culture plates (Falcon), containing LB freezing buffer (LB supplemented with 36 mM K_2HPO_4, 13.2 mM KH_2PO_4, 1.7 mM sodium citrate, 0.4 mM $MgSO_4$, 6.8 mM $(NH_4)_2SO_4$, 4.4% (v/v) glycerol) with a disposable 96 needle inoculation tool. After growth overnight at 37°C, the cultures in microtiter plates are sealed using self-adhesive plastic seals (USA Scientific Plastics) and stored at -80°C. For PCR, cultures are diluted 1:100 in 10 mM Tris-HCl, pH 8.5 and 10 µl of the dilution per 50 µl PCR reaction is used. To fill 96-well plates with liquid a Qfill2 machine (Genetix, Christchurch, GB) is used. To pick colonies grown on plates into a 96- or 384-well format, or in an array on a membrane, a Qpix robot (Genetix) is used. Membranes are put onto an agarose plate and colonies grew overnight. Cells are lysed on the membrane, and DNA is fixed on the support as described by Nizetic *et al.* (1991).

B. Two-dimensional protein gel electrophoresis

For protein gels, 0.5 g frozen, pulverized tissue is suspended on ice in 200 µl SDS boiling buffer (5% SDS, 5% β-mercaptoethanol, 10% glycerol, 60 mM Tris-HCl, pH 6.8). The sample is spun down at 4°C at 10000g in a tabletop centrifuge and the supernatant is transferred into a new tube and kept on ice. Protein concentration is determined as described before (section 8A) using the Bradford reagent. A total amount of 100 µg protein is loaded. Two-dimensional (2D) gel electrophoresis is performed according to the method of O'Farrell (1975) by Kendrick Labs, Inc. (Madison, WI) as follows: Isoelectric focusing (IEF) is carried out in glass tubes of inner diameter 2.0 mm, using 2.0% pH 4 - 8 ampholines (BDH from Hoefer Scientific Instruments, San Francisco, CA) for 9600 volt-hrs. One µg of an IEF internal standard, tropomyosin protein, with lower spot of Mr 33,000 and pI 5.2 is added to the samples. An arrow on the stained 2D gels indicates this standard. After equilibration for 10 min in Buffer 'O' (10% glycerol, 50 mM dithiothreitol, 2.3% SDS and 63 mM Tris-HCl, pH 6.8) the tube gels are sealed to the top of stacking gels which are on top of 10% acrylamide slab gels (0.75 mm thick) and SDS slab gel electrophoresis is carried out for about

4 hrs at 12.5 mA. For silver staining, the slab gel is fixed in a solution of 10% acetic acid/ 50% methanol overnight. The following proteins (Sigma Chemical Corp.) are added as molecular weight standards to the agarose which sealed the tube gel to the slab gel: myosin (M_r = 220.000), phosphorylase A (M_r = 94,000), catalase (M_r = 60,000), actin (M_r = 43,000) carbonic anhydrase (M_r = 29,000) and lysozyme (M_r = 14,000). These standards appear as horizontal lines on the Coomassie Brilliant Blue R-250 or silver stained acrylamide slab gels. For Coomassie Brilliant Blue staining of the protein bands, the polyacrylamide gel is soaked in the staining solution (0.1% w/v Coomassie Blue, 16% v/v acetic acid, 42% v/v methanol) for several hours. The gel is destained in aqueous acetic acid (12% v/v isopropanol, 16% v/v acetic acid.

For silver staining (Switzer *et al.*, 1979), the gel is fixed at room temperature successively in 20% (w/v) trichloroacetic acid for one hour, twice for 30 min in 40% (v/v) ethanol, 10% (v/v) actic acid and twice in water for 20 min. The gel is soaked for 30 min in a 10% (w/v) glutaraldehyde solution, followed by 3 washes in water (20 min each). The proteins are stained for 30 min in silver diamine solution (freshly made up 0.26% (w/v) NaOH, 1.8% (w/v) ammonia, 3% (w/v) silver nitrate), washed three times in water and developed for 10 min in developing solution (0.05% (w/v) citric acid, 0.02% (v/v) formaldehyde) then transferred into stop solution (40% (v/v) ethanol, 10% (v/v) acetic acid).

The stained gels are dried between sheets of cellophane.

C. DNA microarray

To amplify the *Arabidopsis* cDNA inserts cloned into M13 derived plasmids, a PCR on colonies is performed. Bacteria are subcultured in 96 well format plates overnight in LB freezing media (supplemented with Amp). Aliquots of the cultures are diluted 1:100 in 10 mM Tris-HCl, pH 8.5 and 10 µl of the dilutions are used per 50 µl PCR. The PCR mix contained per 50 µl reaction 5 µl AmpliTaq buffer (Perkin Elmer), 10 µl dNTP mix (20 mM each), 2 times 5 µl modified primers (M13 forward: 5'amino tgtaaaacgacggccagt - SEQ ID NO:7, M13 reverse: 5'amino ggaaacagctatgaccat - SEQ ID NO:8, 10 µM each), 1 µl AmpliTaq Gold (Perkin Elmer). PCR thermocycling is performed in a Perkin Elmer 9700 PCR machine as follows: 10 min at 95°C, 40 cycles of 30 sec at 95°C, 30 sec at 51°C, 2 min extension at 72°C, followed by a 10 min extension at 72°C. 5 µl of all PCR reactions are run

on a 1.2% agarose gel (1 x TBE, 6 V/cm). After further purification through QIAquick-96 microfiltration columns (Qiagen) and lyophilization, PCR products are resuspended in 10 µl of 3 x SSC and spotted onto silane-coated glass slides (Synteny, Inc., Fremont, CA). The DNA is rendered single stranded by heat or alkali treatment.

For probe preparation, single strand reverse transcription from an oligo-dT primer is performed in presence of Cy3-dCTP or Cy5-dCTP (Amersham, Arlington Heights, IL) using 600 ng polyA RNA per sample. The reverse transcription reaction is performed in a 25 µl volume with 2 µg oligo(dT) 21-mer, 500 µM each of dATP, dGTP and dTTP, 280 µM dCTP, 40 µM of Cy3 dCTP or Cy5 dCTP, 40 units RNAsin (Promega) and 200 units SuperScript II reverse transcriptase (Life Technologies) in 1 x SuperScript first strand buffer. Three specific control polyA RNAs (YCFR06, YCFR22 and YCFR23) from non-coding yeast genomic DNA are added into the reverse transcription reaction at 0.006 ng, 0.06 ng and 0.6 ng, respectively (ratios of the control RNA to polyA RNA are 1:100,000, 1:10,000, and 1:1000 (w/w), respectively). After incubation at 37°C for 2 h, the reactions of two samples (Cy3 and Cy5) are combined and treated with 5 µl of 0.5 M sodium hydroxide and 5 µl of 10 mM EDTA for 10 min at 65°C to stop the reaction and degrade the RNA. Probes are purified using two successive Chroma Spin 30 gel filtration spin columns (Clontech) and lyophilized. Probes are resuspended in 20 µl hybridization buffer (5 x SSC, 0.2% SDS) and applied to the microarray (Schena *et al.*, 1996). Hybridization is carried out at 60°C for at least 12 hours (Synteni). The slide is rinsed for 5 min each in 5 x SSC, 0.1% SDS and in 0.2 x SSC, 1% SDS at room temperature. Two-channel, confocal laser microscopes are used to scan the fluorescence emission after excitation at 532 nm and 633 nm (Shalon *et al.*, 1996). Local background in the surrounding area of each signal is automatically subtracted and for each experiment, a normalization over all signals is carried out. The expression quantification is scaled from -100 to +100 based on the three control RNAs and on consideration of various complex tissue and standard elements (further information can be found at http://www.synteni.com/client/controls/index.html). The indicated detection limit of the DNA microarray is 1 molecule per 100,000 RNA species (which equals 1 copy per cell) and the differential sensitivity can be as good as twofold. The raw data are fed into GemTools software for preliminary characterization. Extensive comparisons between gene groups are

done in Excel (Microsoft Corp., Redmond, WA), using the S Plus dist function for similarity calculations (Version 3.3, Mathsoft, Inc., Seattle, WA).

D. Diversity of conditions used to describe the transcriptome of *Arabidopsis thaliana* during SAR.

18 inducing conditions are as follows: Two-weeks old plants grown on herbicide (0.9 μM hydantocidin (1)), or adenylosuccinate synthetase-antisensed plants (2); leaf tissue from four-week old mutants *cim6* (3), *cim7* (4), *cim11* (5a); for comparison of reproducibility, sample *cim11* was also inversely labeled with the flourochroms (5b); *NIM1* overexpressing line 8C (6); a NahG plant line (7) ; tissue from crosses of NahG to *cim6* (8) and *cim11* (9); systemic, secondary tissue (10) and primary, infected tissue (11) of NahG plants, locally inoculated with incompatible pathogen *Pseudomonas syringae* pv. tomato (Pst) DC3000 *avrRpm1* at 106 cfu/ml harvested 44 hours after inoculation; tissue harvested after 4 hours (12) or 48 hours (13) after treatment with 0.3 mM BTH; tissue harvested 48 hours after inoculation with a suspension of 105 spores/ ml of the compatible isolate of *Peronospora parasitica* pv. Emwa (14) and of infected primary (15) and systemic secondary (16) tissue of wild-type plants harvested 44 hours after inoculation with Pst DC3000. *avrRpm1*; and systemic tissue of *niml*-4 plants (Ryals *et al.*, 1997) treated with Pst DC3000 *avrRpm1* as above (17); tissue harvested 48 hours after inoculation with a suspension of 105 spores/ ml of the incompatible isolate of *Peronospora parasitica* pv. Noco2 (18).

EXPERIMENTAL - RESULTS

1. PR-1/Luciferase Line Construction And Characterization

A. PR-1 phylogeny

The expression of the *PR-1* gene is the most reliable marker for the onset of SAR in *Arabidopsis* (Uknes *et al.*, 1992). The *PR-1* gene encodes a small (preprotein: 17677 Da, 161 aa; cleaved: 14880 Da, 135 aa), acidic (pI of approximately 4.0) apoplastic protein. Although the function of the PR-1 protein remains unknown, several studies have shown that PR-1 might play a direct role in conferring resistance to fungal pathogens. PR-1 has an

antimicrobial activity *in vitro* and confers resistance to oomycetes when overexpressed *in planta* (Alexander *et al.*, 1993). Because of the central role of the PR-1 gene for this study, a similarity BLAST (Altschul *et al.*, 1990; Gish and States, 1993) search against the *Arabidopsis* PR-1 protein sequence was performed in Genbank (www.ncbi.nlm.nih.gov/htbin-post/Entrez, default parameters). PR-1 protein homologs were found not only in di- and monocotyledonous plants, but also in mammals and insects: PR-1 belongs to a family of cystein-rich secretory proteins (CRISPs) that groups mammalian SCP/TPX1 (sperm coating glycoprotein/testis specific protein) insect AG3/AG5 (venom allergen), fungal SC7/SC14 (*Schizophyllum commune*), and plant PR-1 proteins together. The sequence identities of the homologs range between 30% and 80%. While the function of the human PR-1 homolog is less well understood, in insects, PR-1-like proteins make up a major venom allergen. The close sequence homologies of the PR-1 protein to secreted cystein-rich proteins suggest a possible function of PR-1 in defense similar to other small antimicrobial peptides, such as plant defensins or thionins. This finding is consistent with the previously observed effects of PR-1 on fungal pathogens.

B. Cloning and characterization of the transgenic *PR-1*/luciferase line

A 4258 bp upstream *Xho*I fragment of the *Arabidopsis PR-1* promoter was cloned in front of the luciferase coding region in the binary vector pCIB200 (*see*, U.S. Patent Application No. 08/992,801 and U.S. Patent No. 6,031,153). Stable transformants in the ecotype *Col-0* were selected on GM plates containing kanamycin. Segregation was analyzed in the T2 populations of selfed primary transformants (T1). Several lines that segregated in a 3:1 ratio were chosen for preliminary *in vivo* counting of background (non-induced) luciferase activity and chemically induced activity.

One line, called 6E line, was chosen for further characterization. In a F2 population of an outcross to untransformed *Col-0* plants, 147 out of 203 plants survived on selection for kanamycin resistance. The χ^2 analysis for a 3:1 segregation ratio calculates to $\chi^2 = 0.724$, $P < 0.4$. Southern analysis using either the luciferase gene or the right border of the T-DNA (RB) as a probe showed that only one insert was integrated into the genome. In particular, the timing of the expression of the transgene in the 6E line was expected to match the expression pattern of the endogenous *PR-1* gene following both biological and chemical inductions.

I. Chemical induction of the PR-1/luciferase gene

For determination of chemical inducibility of luciferase activity, *PR-1*/luciferase plants were sprayed with three chemical activators of SAR: SA (5 mM), INA (375 μM) or BTH (375 μM or 5 mM) and luciferase activity was determined every 24 hours during a period of four days. For each measurement, six samples consisting of six leaves each were harvested. INA and BTH treatment at the standard concentration of 375 μM caused an induction of luciferase activity of more than 2000 fold within 48 hours and this level was maintained for at least two more days. In contrast, treatment with 5 mM SA lead to a 5000 fold induction of luciferase activity within 24 hours which was followed by a decline, possibly due to degradation or conjugation of SA. Treatment with 5 mM BTH caused an induction similar to SA but no pronounced decline of activity was observed within the first four days. RNA was extracted from the same tissue samples as used for determination of luciferase activity and *PR-1* mRNA levels were examined by Northern blot analysis. Luciferase mRNA is very unstable in plants (no plant polyadenylation signal was added, as the firefly luciferase gene possesses its own terminator sequence which is active in plants; Ow *et al.*, 1986) and can not readily be detected in Northern blot analysis. Because of the rapid turnover of the luciferase enzyme, its activity reflects roughly its transcription rate. Patterns of *PR-1* mRNA and luciferase activity were comparable, showing that luciferase activity reflects *PR-1* transcription in these plants. To test whether the onset of SAR gene expression could be monitored *in vivo*, *PR-1*/luciferase plants were treated with 375 μM BTH and sprayed with luciferin 2 days after treatment. Light emission of seedlings could be detected *in vivo* within 2 minutes using the photon-counting device.

II. Biological induction of the *PR-1*/luciferase line

To determine the systemic activation of *PR-1*/luciferase subsequent to pathogen inoculation, *PR-1*/luciferase plants were sprayed at 24 and 12 hours before the pathogen treatment with 7.5 mM luciferin to inactivate luciferase (Millar *et al.*, 1992) and to reduce background induction. At 0 hours, approximately 50% of the area of fully developed leaves were infiltrated with either water or with a spore solution of *P. parasitica Emwa* (Table 5).

The incompatible interaction triggered a more than 150 fold systemic induction of the luciferase activity within three days. This induction could also be followed *in vivo*.

Table 5: Induction of the *PR-1*/luciferase transgene after treatment with the avirulent pathogen *P. parasitica Emwa* (Emwa). Luciferase activity was inactivated 24 and 12 hours before the experiment by luciferin treatment. Values represent relative inductions of luciferase *in vitro* activity compared to water treated plants (set to 1). Each time point consists of 20 leaves infiltrated with either water or a spore solution of *P. parasitica Emwa* (10^5 spores per ml) derived from 10 different plants.

Treatment	0 hours	12 hours	24 hours	72 hours
H_2O	1+/-1	1+/-0	1+/-0	1+/-0
Emwa	1+/-0	3+/-1	11+/-5	162+/-68

2. Mutagenesis And Mutant Screen

In order to identify mutants that constitutively express the *PR-1*/luciferase gene, the 6E line was submitted to EMS mutagenesis. Mutant screens in *A. thaliana* are usually performed in the M2 generation. In the M1 generation, the mutations are heterozygous and the plants are chimeric, since in *A. thaliana* seeds, at the time of the chemical mutagenesis, 12 cells represent the origin of the vegetative parts of the M1 plants. The size of the M1 population is dictated by the cost of the mutagenesis on the one hand and the size of the genome, and the mutation rate of the mutagen (mutations per genome and generation) on the other hand. In mature *A. thaliana* seeds, two diploid cells contribute to the formation of the inflorescence (Redei, 1975) and thus to the germ line (so called GECN = genetically effective cell number). We used 8400 M1 plants, corresponding to 4 x 8400 genome equivalents, pooled in 168 lots. This corresponds to roughly 33600 x 80 x 4 = 10,752,000 point mutations in the M1 population, based on the empirical assumption that about 80 homozygous point mutations per M2 plant are generated at our mutation rate (M-value = 0.147; Mednik, 1988; Haughn and Somerville, 1987). This number has to be compared to 21,000 to 25,000 genes of the *A. thaliana* genome (Goodman *et al.*, 1995).

The required size of the M2 population can be derived from the size of the M1 to maximally exploit the genetic potential of the M1. One has to consider that M2 seeds were

pooled from approximately 50 independent M1 lines, and that only half of the M2 siblings of a M1 plant can potentially carry one certain mutation (GECN = 2). For the detection of a recessive mutation (p = 0.25 in the M2), we expect therefore one plant out of 400 progenies (0.5 x 0.25 x 1/50) to show the mutational phenotype. To detect this plant with a high probability (P = 0.95), at least n plants per lot need to be screened, with

$n = \ln(1-P)/\ln(1-1/E)$

n: number of individual plants to be screened in each lot to detect any given mutation

E: expectation value, here 400

n = 1200.

We screened several M2 lots under different culture conditions since it was known that some SAR mutants are conditional (Dietrich *et al.*, 1994). Therefore, an average of 1500 plants per lot or 250,000 plants in total, were screened for spontaneous expression of the *PR-1*/luciferase gene. This represents a coverage of the M1 gene pool in the M2 generation of

$P = 1 - (f)^n$ (Redei and Koncz, 1992)

with

P: probability to detect a recessive homozygous mutation in the M2

n: number of M2 plants screened per M1 plant = 250000 / (168 x 50) = 29.7

f: theoretical fraction of M2 plants that do not show a mutation present in one of the proposed effective two germ cells of a M1 plant = 1 - 1/8

$P = 1 - 0.875^{29.7} = 98.1\%$

Thus, the probability to identify a given M1 mutation in the M2 population approached saturation.

The frequency of putative mutants did however not vary with the conditions, and was approximately 2.4×10^{-3}. 160 M1 pools contained at least one plant that constitutively expressed the *PR-1*/luciferase gene.

In total, 603 putative mutants were identified in this *in vivo* screen. Almost all of them were confirmed by the *in vitro* luciferase analysis in the M2 or their progeny. The phenotypes were therefore considered to be caused by genetic mutations. We expected a multigenic regulation of *PR-1* gene expression and, thus, a high number of mutants. However, the number of mutants was further increased by the redundancy of mutant identifications in the M2, which provided several fold coverage of the M1 gene pool (M2 saturation of the M1) and the fact that several independent mutations per gene (alleles) might have been identified (M1 saturation of the genome). To find out how many independent genes are actually involved in the SAR cascade, all 603 mutants would have to be mapped or crossed to each other.

Considering the frequency of *PR-1*-overexpressing plants in an M2 after EMS mutagenesis, T-DNA insertion mutagenesis might also yield mutants with constitutive *PR-1*/luciferase expression, within a reasonable population size. Using T-DNA insertions to mutagenize the genome has the disadvantage of being labor-intensive and having a low mutation rate (1-2 inserts per genome, compared to over 80 point mutations per genome by EMS), but has the potential advantage of easy cloning of the mutated, "tagged' gene, by plasmid rescue or other techniques. Because of our finding from the EMS screen, that most of the *cim* mutants were dominant, or codominant (see below, section 3B), we screened the T1 population (primary, hemizygous transformants) for constitutive *PR-1*/luciferase expression. 10,000 T1 lines were screened for *in vivo PR-1*/luciferase activity, and 80 of them were retested in the T2 generation. Strong luciferase activity was confirmed in 7 T2 populations (see section 4E, table 15). This low rate of confirmation in the T2 generation was caused by the low cut-off of luciferase activity that was used in the T1 as a criteria for selection in order to also find codominant mutants. Therefore, more false positive plants were retained than in the EMS screen.

3. Characterization Of *cim* Mutants

A. Histological characterization of the *cim* mutants

The goal of this study was to identify SAR activated mutants that do not show spontaneous cell death, according to the definition of *cim* mutants. In a secondary screen, the 603 mutants were therefore subjected to a Trypan Blue lesion staining in the M2 and the M3

generation and examined both macroscopically and microscopically for cell death. More than 90 mutants did not show macroscopic patches of cell death, but only 16 did not have any cell death under our growth conditions, as revealed by microscopy after staining. Most of the other mutants developed necrotic lesions at some stage of their life cycle, mostly in the leaf tips.

One of the mutants with spontaneous cell death, designated mutant 779, was included in all the following experiments as a control. Mutant 779 displayed patches of autofluorescence and callose that normally accompany HR-like cell death. No callose was detected in the 16 *cim* mutants. Although free of lesions, pleiotropic phenotypic alterations in the 16 *cim* mutants were not separated from the mutation that caused constitutive *PR* gene expression by three backcrosses. In general, *cim* mutants have a prolonged life cycle, a delayed flowering time (one to four weeks later than wild-*type Col-0*) and they set fewer seeds (approximately one third of *Col-0*). Some mutants also showed reduced germination. Leaf morphology varied from long, often curly leaves (*cim*205, *cim*716), to extremely small, round leaves (*cim*677, *cim*810). Mutant *cim*677 showed a bright green leaf pigmentation, other *cim* mutants (*cim*713, *cim*810) had dark-green leaves. However, normal leaf morphology was also found, albeit mostly in the weaker mutants, *cim*328 and *cim*658 (weakness based on *PR-1* gene expression and SA content, see below) as well as in the mutant *cim*713 that differed from wild-type only in size.

B. Genetic characterization of the *cim* mutants

All 16 mutants originated from different seed pools and were therefore considered independent mutations. All mutants were backcrossed at least three times to the *PR-1*/luciferase parental line.

Table 6: Genetic characteristics of the 16 *cim* mutants identified in the *PR-1*/luciferase EMS mutant screen.

	16 *cims* [1]	M1 pool [2]	Luc in F1 [3]	F2 (+ : -) [4]	χ^2 (P) [5]	Map position
Mutants described in more detail	8	36	yes	57 : 25	1.32 (P>0.2)	chr. 2 [9]
	205	15	yes	40 : 20	2.22 (P>0.1)	chr. 1
	328	2	yes	n. d[7].		
	658	93	yes	58 : 16	0.45 (P>0.5)	
	677	98	no	33 : 99	0 (P>0.95)	
	695	107	yes	31 : 11	0.03 (P>0.95)	chr. 5
	713	115	yes	43 : 12	0.30 (P>0.8)	chr. 1
	716	124	yes	34 : 21	5.10 (P>0.01)[8]	
	[779]	168	yes	47 : 60	55.1 [8]	
	810	164	yes	33 : 19	3.69 (P>0.05)	
	824	155	yes	35 : 11	0.03 (P>0.95)	
Not further characterized mutants	2	17	no	n.d.		
	11 [6]	36	yes	87 : 35	0.89 (P>0.4)	
	81	8	yes	n.d.		
	367	16	yes	n.d.		
	671	91	yes	28 : 22	9.63 (P>0.01) [8]	
	714	121	yes	15 : 5	0 (P>0.95)	
	741	132	yes	22 : 18	8.53 (P>0.01)[8]	

1. identification number of *cim* mutant
2. M1 lot of mutant origin
3. luciferase activity in F1 of backcrosses to the 6E line at least five times above background; due to incomplete penetrance of the mutant phenotype, a varying percentage of plants in the F1 populations expressed the *PR-1*/luciferase gene
4. segregation ratios of luciferase expressing to not-expressing plants in the F2 generation
5. χ^2 and probability of the observed difference to the expected 1:3 segregation ratio
6. *cim*11 originated from the same M1 seed lot as *cim*8, they may be identical
7. n.d.: not determined
8. hypothesis of a 3:1 segregation rejected
9. chr.: chromosome

Selfed progeny of all mutants stably expressed *PR-1*/luciferase. The presence of the mutations could normally be followed in all mutants by *in vivo* monitoring of luciferase activity in the F1 generation.

To analyze the segregation ratios of the mutations, F2 populations of backcrosses, containing 20 to 100 plants were screened for constitutive luciferase activity and the resulting data were subjected to $\chi 2$ analysis (Table 6). The expression of the reporter gene in the F1,

confirmed in random samples by Northern blot analysis for endogenous *PR-1* expression, indicated that in all but two cases (*cim*2, *cim*677) the mutant phenotype was dominant. However, the analysis of the F2 segregation ratios suggested that many of these mutations were not fully penetrant. In addition, we can not exclude the possibility that in some cases (*cim*716) two dominant genes are required to cause the observed phenotype ($\chi 2$ for a 9:7 segregation ratio = 0.69, P<0.4). In the case of *cim*713, the morphological changes were inherited in a recessive manner, while the closely linked constitutive *PR-1*/luciferase expression was codominant, with varying expression of the phenotype in the heterozygous plants. In cases where F2 segregation ratios were normal, we mapped the mutations to the responsible loci. Populations of usually 50 to 80 F2 plants (mutant crossed to ecotype *Ler*) with preselected phenotype were used to look for linkage between the *cim* phenotype and genetic markers. About 30 SSLP and CAPS markers that were evenly distributed throughout the genome, were chosen to find a primary linkage. If available, closer markers were chosen to further define this map position. *cim*713 was placed on the genetic map of *Arabidopsis thaliana* on chromosome 1 between markers mi291a and markers nga280 (see below, section 4B). *cim*205 is also located on chromosome 1, between markers nga280 (20 recombinants in 116 analyzes meiosis) and m185 (19 recombinants in 148 meiosis). *cim*8 is located on chromosome 2, between markers ve017 (16 recombinants in 148 meiosis) and nga168 (9 recombinant in 122 meiosis). *cim*695 lies on chromosome 5 between markers DFR (22 recombinants in 106 meiosis and LFY (17 recombinant in 110 meiosis). The map positions of the mutations on chromosome 1 and 2 do not match the map position of known mutations in genes encoding functions in disease resistance and/or SAR. Mutant *cim*695 is in a region of chromosome 5 termed MRC-J, which contains a number of *R* gene homologs (Botella *et al.*, 1997; Holub and Beynon, 1997). *cim*205 and *cim*713 map close to but distinct from *cpr6* (Clarke *et al.*, 1998).

C. Biochemical characterization of the *cim* mutants

I. *cim* mutants accumulate salicylic acid

One characteristic of SAR is the accumulation of SA, which is not only sufficient to induce *PR* gene expression and resistance, but also required, as shown with the SA-degrading NahG lines. Free and total (= free and conjugated) SA content was measured in 90 mutants.

Although we expected to find some mutants downstream of SA, for example possible gain-of-function *nim1/npr1* alleles, all mutants (with the exception of *cim*658) accumulated 3- to 15-fold more SA than untreated wild-type plants (Table 7). A control treatment with a virulent *Erysiphe cichoracearum* pathogen caused a 7-fold increase in total SA content after three days of infection.

Table 7: Salicylic acid content is increased in most *cim* mutants. Total salicylic acid content was determined by HPLC separation of an organic tissue extraction. Results (in ng SA per mg tissue fresh weight) are mean values and standard deviations of three independent measurements. As a comparison, SA content was measured in wild-type tissue infected with *Erysiphe cichoracearum*, harvested 3 days after inoculation.

mutant	total SA (ng/mg fresh weight)
wild-type	296 +/- 25
wild-type + *Erysiphe*	2030 +/- 890
*cim*8	1958 +/- 835
*cim*205	1657 +/- 436
*cim*328	899 +/- 16
*cim*658	294 +/- 21
*cim*677	4154 +/- 211
*cim*695	2256 +/- 223
*cim*713	1500 +/- 78
*cim*716	1350 +/- 267
*cim*810	3415 +/- 331
*cim*824	2190 +/- 491

On average, among the 90 mutants for which SA content was measured, lesion mimics have higher levels of SA than *cim* mutants, but the two distributions are overlapping: *cim* mutants can have more SA than lesion mimic mutants. *cim* mutants are hence not simply a weak subclass of lesion mimic mutants, but ought to be considered as a distinct class of mutants. Free SA content was about tenfold less than total SA and was always correlated to the total SA content thus excluding from our collection mutations in the regulation of this equilibrium or in the degradation/conjugation of SA. Based on SA content (and *PR* gene expression, see below), mutants can be classified into strong *cim* mutants (e.g. *cim*8, *cim*205, *cim*677, *cim*695, *cim*810, *cim*824) and weak *cim* mutants (e.g. *cim*328, *cim*658).

II. Other phenylpropanoids in *cim* mutants

To determine if the elevated SA levels in the cim mutants were the result of altered regulation at a specific step in the SA biosynthesis, attempts were made to measure levels of SA precursor molecules. We were however not able to detect *t*-CA or benzoic acid in plant extracts, because of difficulties in detection. Unlike SA, these compounds emit no detectable fluorescence (λ_{max} at 230 nm).

In a cooperation with Prof. Dr. Strack (Halle, Germany), the phenylpropanoid content in the ten *cim* mutants was analyzed in more detail, using two different organic extractions (80% methanol, or dichloromethane) and a basic cell wall extraction procedure, followed by a separation on a C-18 reverse phase HPLC. With minor, unidentified exceptions, no variations were detected in the UV spectra of the mutants compared to wild-type.

III. *cim* mutants can accumulate low levels of camalexin

Plants under pathogen attack accumulate antimicrobial molecules, called phytoalexins. The major phytoalexin in *Arabidopsis* is camalexin, derived from a tryptophan precursor. Camalexin can be visualized under UV light after TLC separation of methanol plant extracts. Both a synthetic standard and tissue harvested after *Pseudomonas* infection were used as controls. The R_f for synthetic camalexin was a bit lower (0.77) than the R_f for camalexin in complex mixtures (0.78). Mutants *cim*328, *cim*677, *cim*716, 779 and *cim*810 accumulated between 1 μg camalexin/cm^2 (based on the synthetic camalexin standard) and the amount that accumulated in *Pseudomonas*-infected tissue after 3 days (literature value: 1.2 μg camalexin/cm^2; Glazebrook and Ausubel, 1994). Mutant *cim*824 accumulated less than 1 μg/cm^2, and *cim*8, *cim*658, *cim*695 and *cim*713 did not show substantial amounts of camalexin. A quantification of the camalexin by HPLC of samples isolated from TLC plates was not successful, possibly because the amount of camalexin was too small for the available detection system.

D. *cim* mutants exhibit resistance to pathogenic microorganisms

In order to show that SAR (or LAR) is constitutively activated in *cim* mutants, resistance to SAR pathogens must be shown.

I. Most *cim* mutants are resistant to *Peronospora parasitica*

The response of the mutants to two isolates of the oomycete parasite *Peronospora parasitica* which are virulent on wild-type *A. thaliana Col-0* (Holub *et al.*, 1994) was tested. Infection was scored in adult leaves eight days after inoculation with *P. parasitica Noco2*. While *Col-0* wild-type plants were covered with hyphae and displayed cell death along the fungal structures, no fungal structures, chlorosis or spontaneous macroscopic necrosis was observed in the "strong" *cim* mutants with high levels of *PR-1* gene expression and SA eight days after infection. Only the two "weak" mutants *cim*328 and *cim*658 allowed fungal hyphal growth and slight sporulation. Trypan Blue staining for hyphal growth and cell death revealed, however, that in some mutants trailing necrosis or HR like pits (Holub *et al.*, 1994) occurred around the hyphal penetration sites (e.g. *cim*695, *cim*716). This phenomenon was not correlated to SA content or *PR* gene expression. A second compatible *P. parasitica Emco5* (Holub and Beynon, 1997; McDowell *et al.*, 1998) was applied to younger plants in a cotyledon assay, because infection of wild-type plants is stronger at this earlier stage. Resistance to *P. parasitica Emco5* was similar to that observed with the isolate *Noco2*. The results suggest that the observed resistance is not age-dependent or an isolate-specific reaction.

II. A subset of *cim* mutants is resistant to *Erysiphe cichoracearum*

To determine if the disease resistance seen in the *cim* mutants was specific to *Peronospora* or more broad-spectrum, the mutants where challenged with a second fungal pathogen. *Erysiphe cichoracearum* is virulent on most *A. thaliana* ecotypes, including *Col-0* (Adam and Somerville, 1996). To have a negative control for pathogen infection, it was first shown, that *Col-0* can be completely protected from *Erysiphe* infection by BTH treatment (0.3 mM). A disease rating system between 1 (resistant) and 3 (susceptible) was used to quantify macroscopic symptoms. Interestingly, this assay revealed a differential response among the *cim* mutants. Some *cim* mutants (e.g. *cim*328, *cim*810) are completely resistant to *E.*

cichoracearum, and others completely susceptible. The resistance did not correlate with the strength of *PR-1* gene expression or SA content. The two strongest mutants *cim*677 and *cim*810 were resistant, but *cim*328, with low *PR-1* gene expression and SA accumulation, also displayed an almost complete resistance (disease rating 1.01).

III. Some *cim* mutants exhibit a good resistance to virulent *Pseudomonas syringae* pathovars

To check the *cim* mutants for resistance to bacterial pathogens, they were inoculated with several different virulent *Pseudomonas syringae* strains. Interpretation of the results of these experiments was often hampered by the non-wild-type morphology and developmental stage of the *cim* mutants. It became clear, however, that resistance to *Pseudomonas* spp. was in many mutants not as clear-cut as resistance to *Peronospora* and *Erysiphe*. Differences in resistance to the aggressive pathogen *Pseudomonas syringae* pv. *syringae* DC3000 were small among the mutants. Therefore the less virulent *strain P. syringae* pv. *maculicola* ES4326 was chosen to better illustrate the spectrum of resistance to *P. syringae* among these mutants. Mutant *cim*677, *cim*695, *cim*713, and *cim*810 exhibited a bacterial proliferation reduced more than 10-fold compared to wild-type (Table 8). For mutants *cim*205 and *cim*716, the bacterial titer five days after inoculation was 2-fold lower than in wild-type. While mutants *cim*8 and *cim*824 are both in the class of "strong" mutants, they were at least as susceptible to this *P. syringae* isolate as wild-type (Table 8).

The response to avirulent *Pseudomonas* strains also differed among the *cim* mutants. Several *cim* mutants (*cim*205, *cim*328, *cim*716, *cim*824) displayed an HR identical in timing and appearance to wild-type. Other *cim* mutants (*cim*713, *cim*810) did not show the typical HR symptoms or HR was delayed. In *cim*677, the HR occurred 2 to 3 hours earlier than in wild-type, in which the HR was visible 8 to 9 hours after inoculation. This mutant showed increased sensitivty to virulent *Pseudomonas* strains, although the injection of $MgCl_2$ did not have any visible effect on the plant.

Table 8: Resistance of *cim* mutants to *Pseudomonas syringae* pv. ES4326. Bacterial colony forming units (cfu) were counted at three and five days after infiltration from four independent experiments, each containing four leaf punches per mutant per time point. While in most *cims* bacterial growth is significantly limited, in some mutants (*cim8, cim824*) proliferation is not reduced. 6E: wild-type, B: wild-type pretreated with 0.3 M BTH 2 days prior to *Pseudomonas* infection.

mutant	mean cfu, 3dpi [2]	*t* value (P), 3 dpi[3]	mean cfu, 5 dpi [2]	*t* value (P), 5 dpi[3]
6E	965000		2140000	
B	36500	5.82 (P>0.995)	38000	1.74 (P>0.9)
8	1752500	0.53 (P>0.65)	12550000	3.37 (P>0.99)
205	333500	2.53 (P>0.975)	995000	0.60 (P>0.7)
328 [1]	151500	0.94 (P>0.8)	68000	3.23 (P>0.99)
677	105000	5.40 (P>0.995)	35500	1.74 (P>0.9)
695	467500	1.33 (P>0.85)	69000	1.71 (P>0.9)
713 [1]	25650	1.74 (P>0.9)	14200	4.66 (P>0.995)
716	82500	5.47 (P>0.995)	840000	1.04 (P>0.8)
810	618666	1.13 (P>0.85)	273333	1.50 (P>0.9)
824	730000	0.56 (P>0.7)	3350000	0.58 (P>0.7)

1. Mutant *cim328* and *cim713* were tested in a separate experiment, the cfu values for 6E wild-type were at 3 dpi: 327000 and at 5 dpi: 362500
2. Bacterial viable count, expressed as cfu (colony forming units per 4 leaf discs), calculated from four independent repetitions, dpi: days after inoculation
3. Differences of mean values of 6E wild-type and mutants, expressed as *t* values

E. Molecular characterization of the cim mutants

I. Northern blot analysis

To further characterize the mutants at molecular level, gene expression for known marker genes of various pathways was analyzed. The results may provide insight into the signaling pathways that are turned on in *cim* mutants and allow a further understanding of the maintenance phase of SAR.

Genes implicated in oxidative burst and cell death

Expression of genes implicated in regulating the cellular redox state or the oxidative burst was monitored in the *cim* mutants. Superoxides and/or hydrogen peroxides are thought to determine the outcome of plant-pathogen interactions by several means: *i.*) they can reinforce the cell wall, *ii.*) they can act as messengers, and *iii.*) they could be involved in direct killing of the pathogen or the host cell during the HR (Low and Merida, 1996). Sources of the reactive oxygen species (ROS) can be cell-wall bound NADPH oxidases, peroxidases and

lipoxygenases. ROS are degraded in the plant cell by superoxide dismutases (SOD), and successively in the Halliwell-Asada (Ascorbate-Glutathione) cycle, reducing H_2O_2 to H_2O under use of NADPH (Asada, 1994; Halliwell and Gutteridge, 1989). Three families of SOD are known, classified by their cofactor (Bowler *et al.*, 1992). It has been previously shown that Cu/ZnSOD gene expression is suppressed during pathogen infection and oxidative stress (Fodor *et al.*, 1997). In several *cim* mutants (*cim*716, *cim*810, *cim*658), Cu/ZnSOD gene expression is suppressed. As an alternative to reduction of H_2O_2, catalases may dismutate cellular H_2O_2. In *Arabidopsis*, three genes encode for catalases, that are 70 to 72% identical at the nucleotide level (Frugoli *et al.*, 1996). The specific functions of catalase isozymes are still not well understood. The expression of the catalase 2 and 3 genes is not significantly altered in the *cim* mutants, or by pathogen treatment. Peroxidases also degrade H_2O_2 by oxidizing specific substrates. They are responsible for lipid peroxidation, and cell wall cross-linking. In *Arabidopsis*, little is known about the cellular function and regulation of the two peroxidases used as probes in Northern blot analysis, cationic peroxidase and peroxidase C (Jabs *et al.*, 1996). The genes of both peroxidases are induced in the lesion mimic mutant 779, and to some extent by BTH and *Peronospora*, as well as in *cim*824 and *cim*810. *Peroxidase C* is also weakly induced in the other mutants.

Although one of the selection criteria for the *cim* mutants was absence of cell death, expression of several genes known to cause lesion mimic phenotypes when mutated were also checked for changes in expression. The barley *MLO* gene, whose function is unknown, has several homologs in *Arabidopsis* (Büschges *et al.*, 1997). One of the *ests* (205N12T7) was used to quantify the expression of the *MLO* genes in the *cim* mutants. The degree of homology between the different Arabidospsis *MLO* homologs, and their total number in the genome is not known, it is therefore likely that the signal on the Northern (RNA) blot reflects the expression of several genes and hides potential differential gene expression. No significant changes in gene expression were detected. The same is true for the *Arabidopsis* lesion mimic gene, *lsd1* (Dietrich *et al.*, 1997) and for the *Arabidopsis* homolog of the maize lethal leaf spot gene (*lls1, est* 84E8T7) (Gray *et al.*, 1997). A gene putatively implicated in the regulation of cell death, *DAD* (defense against death; Sugimoto *et al.*, 1995) was slightly induced in all mutants compared to wild-type. Its two-hybrid-interactor, *MOM* did not show major differential gene expression among the *cim* mutants.

Genes implicated in pathogen defense responses

While the expression of genes implicated in the regulation of cell death for the most part did not change significantly in the *cim* mutants, genes whose expression correlates with pathogen infection and disease resistance, were induced in *cim* mutants. The intensity of expression of *PR-1*, whose promoter was used to identify the mutants, correlated with the level of luciferase activity in all mutants. In mutants with weaker luciferase activity (*cim*328 and *cim*658), *PR-1* gene induction was weaker, but still stronger than in wild-type. The strongest mutants in terms of both luciferase activity, and disease resistance accumulated the highest amounts of *PR-1* messengerRNA. *cim*677, *cim*824 and the lesion mimic 779 all showed over 40-fold induction of *PR-1* gene expression, compared to wild-type. *PR-1* gene expression is stronger in these mutants than that obtained by BTH or pathogen induction of SAR in wild-type plants. This finding eliminates the possibility that the luciferase activity used to select the mutants was the result of a mutation in the promoter of the transgene, since all mutants also expressed the endogenous *PR-1* gene to high levels. Expression of the two other *Arabidopsis* SAR genes (Uknes *et al.*, 1992), *PR-2* (data not shown) and *PR-5* was also elevated in most of the *cim* mutants, though not to the same degree as *PR-1*. It has been previously shown that the regulation of these genes does not always correlate with *PR-1* gene expression (Reuber *et al.*, 1998). The *PR-5* gene expression was, however, induced in all mutants, with the exception of the weakest mutant, *cim*658.

Thionins and defensins are genes that are induced by pathogen attack. They are not induced during SAR but are regulated by an SA-independent, JA-dependent signaling pathway (Epple *et al.*, 1998; Penninckx *et al.*, 1996). Both an antagonistic interaction and a concomitant induction of *PDF1.2* and *PR-1* expression have been described in the literature (see Maleck and Dietrich, 1999 for review). Interestingly, *thionin2.1* was induced in some *cim* mutants (*cim*658, *cim*677, *cim*716 and mutant 779). *PDF1.2* is known to be induced in several lesion mimics, and was also induced in mutant 799. It was induced in *cim*695, *cim*677 and *cim*824, but is suppressed in other *cim* mutants (*cim*810, *cim*205, and by BTH). Notably, in some *cim* mutants, the expression level of *PDF1.2* did not correlate with the resistance to *Erysiphe cichoracearum*, as reported previously for some *cpr* mutants (Bowling *et al.*, 1997): For example, the lesion mimic mutant 779 displayed strong *PDF1.2* gene expression but was

susceptible to the fungal pathogen, while *cim*328 showed *PDF1.2* gene expression at uninduced wild-type levels, but almost complete resistance.

Some members of the lipid transfer proteins (*LTP*) gene family have similar induction kinetics as defensins and thionins, and are thought to play a similar role in defense (Molina and Garcia-Olmedo, 1997). In a Northern blot, *LTP* appears to be constitutively expressed, probably due to the cross hybridization of the probe with multiple *LTP* transcripts of the gene family.

Two genes that are involved in the signaling leading to systemic disease resistance have been identified by mutant screens and positional cloning. The *NIM1* gene is modestly SA-inducible (Cao *et al.*, 1997; Ryals *et al.*, 1997) and is induced in the *cim* mutants with the highest SA accumulation (*cim*205, *cim*677, *cim*695, *cim*810, *cim*824). The *NDR1* gene that is induced in incompatible plant-pathogen interactions reflects these weak changes. Interestingly, the previously described induction of the *NDR1* gene by pathogens (Century *et al.*, 1997) was not observed in the pathogen-infected control (tissue harvested eight days after inoculation with *P. parasitica Noco2*). This finding is consistent with the observation that NDR1 function is not required for resistance to *Noco2* (Century *et al.*, 1997).

Genes implicated in secondary metabolism

The expression of the gene of a key enzyme in the shikimate pathway, EPSP synthase (Görlach *et al.*, 1995) was not induced in any of the mutants (data not shown). However, *PAL* genes, which encode the putative rate-limiting enzyme in the general phenylpropanoid pathway leading to SA biosynthesis (Bate *et al.*, 1994) and which are inducible by many biotic and abiotic factors (Wanner *et al.*, 1995) were induced in some mutants (e.g. in *cim*713, *cim*716) and by BTH treatment, but not in others (*cim*695, mutant 779). We used the *PAL1* gene clone as a probe, which is very similar to the *PAL2* gene in *Arabidopsis* (the amino acid sequences are 90% identical), but less similar to the *PAL3* gene, which is usually expressed at very low levels (Wanner *et al.*, 1995). The induction of *PAL* gene transcription did not correlate with the SA content or disease resistance in these mutants. Considering the accumulation of SA in the mutants, and the primary regulation of the phenylpropanoid pathway on the transcriptional level, it is remarkable that not all mutants showed a similar pattern of induced gene expression for key enzymes in this pathway. Preliminary results showed that PAL activity in all *cim* mutants (and a pathogen-infected control) were 1.2 to 1.4

fold higher than in wild-type (except for *cim*658), suggesting a post-transcriptional regulation of PAL activity.

The expression of the chalcone synthase gene (*CHS*; Shirley *et al.*, 1995) paralleled the induction of *PAL* gene expression in all mutants. To our knowledge, in *Arabidopsis*, the induction of chalcone synthase gene expression during SAR has never been studied.

Concomitant with the finding that none of the mutants accumulated high levels of camalexin, phosphoribosyl-anthranilate synthase (*PAT1*) gene expression was only weakly induced in the *cim* mutants. *PAT1* gene expression is tightly correlated to gene expression of the other enzymes in the indole biosynthesis pathway and induced during bacterial pathogen infection (Zhao and Last, 1996). We did not find a strict correlation between accumulation of *PAT1* transcript and camalexin accumulation, although *cim*677, *cim*810, and *cim*824 have all higher levels of camalexin and *PAT1* gene expression.

Hormone-inducible marker genes

The expression of some plant hormone-inducible genes was monitored in *cim* mutants. The *Arabidopsis* vegetative storage protein acid phosphatase gene (*AtVSP*), a marker for jasmonic acid induced gene expression (Berger *et al.*, 1995) was either weakly or not induced.

PR-4 expression is ethylene-inducible. Its expression in the *cim* mutants was about 20-fold weaker than that observed in an ethylene treated control plant and not induced above wild-type level.

A possible exception may be the *Rab18* gene, an example of an ABA-inducible gene (Merlot and Giraudat, 1997). *Rab18* gene expression was induced in *cim*328, *cim*695 and mutant 779.

In conclusion, the induction of SAR in *cim* mutants most likely does not induce or depend on other hormonally regulated pathways as monitored by marker gene expression.

II. Gene expression profiling using a DNA microarray

The gene expression patterns obtained by Northern blot analysis in the different *cim* mutants seemed to be similar to each other, and the observed qualitative differences in pathogen resistance are not easily explainable using these data. To get a more complete description of the transcriptional changes during SAR and in *cim* mutants, a new technology, the DNA microarray, was tested. Because SAR and the *cim* mutants have been quite well

studied in terms of gene expression, they provided a good system to establish this technology (technology verification). At the same time, new marker genes, or gene profiles, might help to unravel the different regulatory mechanisms involved in plant disease resistance (knowledge discovery).

Contrasting other genome-wide gene expression monitoring techniques, such as differential display or cDNA AFLP (Bachem *et al.*, 1996; Diatchenko *et al.*, 1996), the DNA microarray technique can show expression patterns only for cataloged genes (although not necessarily with known function). For *Arabidopsis*, an extensive expressed sequence tag (*est*) collection is available. We used part of the Michigan State University (MSU) non-redundant *est* collection, containing 13,000 clones in 135 96-well plates, as "elements" on the DNA microarray.

On each microarray, a competitive hybridization with two differentially labeled probes was carried out and a signal (per fluorochrome and per element) was integrated and the local background subtracted. The samples prepared for hybridization on the DNA microarray are listed in Table 9.

Table 9: Sample pairs used in competitive hybridization experiments on the DNA microarray

Experiment	cy3 probe	cy5 probe
1	*cim*205	wild-type
2	*cim*328	wild-type
3	*cim*713	wild-type
4	wild-type	*cim*713
5	wild-type + BTH [1]	wild-type
6	NahG	wild-type
7	*cim*205 x NahG	wild-type
8	*cim*713 x NahG	wild-type

1. 0.3 mM BTH were sprayed 2 days prior to harvest

Technology evaluation

We used several controls to evaluate the new DNA microarray technology for its use in plant biology.

(i) element-to-element reproducibility

In order to spot the *ests* on a glass slide, the plasmid inserts were amplified from liquid cultures using a modified PCR primer (5' amino linker). About 90% of the inserts gave sufficient PCR products, but the yield could not be improved by isolating the plasmids prior to

PCR. In addition to the *est* collection, one plate with known and well characterized genes was prepared for spotting. The clones comprised both constitutive genes such as the biotin synthase and protoporphyrin oxidase genes, and pathogen- or stress-induced genes (*PR* genes, oxidative stress genes). To evaluate the element-to-element reproducibility in the same hybridization, all cDNAs of the control plate were spotted three times on the DNA microarray. Table 10 shows the values for 11 genes whose expression was upregulated in the three *cim* mutants compared to wild-type. The standard deviations did not exceed 30% of the mean values and were in many cases much better. Inconsistency in spotting DNAs onto the slide is a more common source of variability than differences in the actual hybridization signal. For example in the hybridization experiment 2 (*cim*328 versus wild-type), the spotting failed twice for the *CHS* gene. No signal was read from two of the three spots. Spotting is still the most delicate process in the microarray fabrication and may fail in a certain percentage of arrays. The precise fail rate of spotting is difficult to estimate because PCR errors also add to the rate of signal failure (seen in Table 10, for the *PAL* and the *LOX* gene), but both together do not exceed 15%. In conclusion, element-to-element variation is at an acceptable level.

Table 10: Element-to-element variability on the DNA microarray. cDNA clones were spotted three times on the array and expression values were compared in three independent hybridization experiments, using as probes RNA from *cim*713, *cim*205 and *cim*328, and wild-type as comparison. Raw data (expression) and mean and standard deviation (mean+/-std) for 11 genes with significant inductions are shown.

Gene	***cim*205**		***cim*328**		***cim*713**	
	expression	mean+/-std	expression	mean+/-std	expression	mean+/-std
PR-1	10.8/9.1/6.2	8.7+/-1.9	10.0/7.5/7.4	8.3+/-1.2	9.0/6.6/2.8	6.1+/-2.6
PR-5	3.6/3.3/3.1	3.3+/-0.2	4.1/3.8/3.8	3.9+/-0.1	6.2/5.9/6.2	6.1+/-0.1
PR-2	4.7/4.7/2.1	3.8+/-1.2	3.4/2.0/3.0	2.8+/-0.6	6.2/6.0/3.3	5.2+/-1.3
PAL[1]	3.5/-/-	3.5	3.4/-/-	3.4	5.2/-/-	5.2
CuZnSOD	1.6/1.4/1.2	1.4+/-0.2	2.2/2.1/2.1	2.1+/-0.0	4.3/2.4/3.4	3.4+/-0.8
LOX[2]	1.6/2.1/-	1.9+/-0.3	1.9/1.7/-	1.8+/-0.1	4.2/3.6/-	3.9+/-0.3
CHS	1.5/1.2/3.7	2.1+/-1.1	3.6/-/-[3]	3.6	3.6/1.2/1.0	1.9+/-1.2
PRXC	2.8/2.6/2.1	2.5+/-0.3	2.1/1.6/1.7	1.8+/-0.2	3.4/2.6/2.3	2.8+/-0.5
NIM1	3.5/2.7/2.2	2.8+/-0.5	2.7/2.4/1.5	2.2+/-0.5	3.2/2.6/2.2	2.7+/-0.4
GLUPER[4]	2.5/2.5/2.5/ 2.4/2.4/2.2	2.4+/-0.1	2.4/2.0/2.0/ 1.9/1.7/1.7	2.0+/-0.2	2.2/1.9/2.0/ 2.1/2.1/1.8	2.0+/-0.1
APX	1.7/1.7/1.6	1.7+/-0.0	1.4/1.3/1.3	1.3+/-0.0	1.8/1.7/1.7	1.7+/-0.0

1 Only one element gave a signal in all three hybridizations, possibly due to a failed PCR reaction
2 One element did not give a signal in any hybridization
3 Two elements did not give a signal in this particular hybridization, thus revealing a spotting error
4 *GLUPER* was spotted 6 times on the DNA microarray

(ii) experiment-to-experiment variability

For each experiment, a pair of probes was prepared. Single strand cDNA was labeled during synthesis with the fluorochromes cyanine cy3 (red) or cy5 (green). These dyes differ only in one double bond and it is likely that they are incorporated with identical efficiency in the first strand cDNA synthesis. The cyanines have very high, but similar molar absorptivity ($\varepsilon > 50000\ cm^{-1}M^{-1}$) and large fluorescence enhancements upon binding to nucleic acids. Cy3 and cy5 have distinct and narrow emission peaks at 532 nm and 633 nm, respectively. To estimate the variation between two experiments, and to estimate the influence of the dyes, we labeled the same pair of RNA twice but reciprocally, i.e. in the first hybridization, we compared the RNA of *cim*713 (cy3) to the RNA of wild-type (cy5), in the second hybridization, we compared wild-type (cy3) to *cim*713 (cy5; experiment 3 and 4 in Table 9). The difference between the value pairs ranges in general between 10% and 30%. These changes do not exceed the element-to-element variability. The linear regression coefficient for the 10 data pairs calculates to $r = 0.83$. The data points from the two experiments can thus be correlated linearly and different hybridization experiments may be compared to each other.

(iii) Comparison of DNA microarray and Northern blot analysis of gene expression

A third validation is a direct comparison between differential gene expression quantified by the DNA microarray technology and by Northern blot analysis (Table 11). With few exceptions, such as the *PR-1* gene expression in *cim*205 and *cim*713, the DNA microarray gives higher absolute values of gene induction than obtained in Northern blot analysis. This is not surprising because the values are normalized on a scale from +100 to -100 which appears to extend the scale (compare section 5B, where the same phenomenon was observed). Apparently, the yeast RNA that was spiked for normalization purposes in the plant samples did not match exactly the abundance of plant mRNA in the cell. Induced expression levels between 3-fold and 6-fold are reported by both technologies to similar levels (e.g. *PR-5*, *PAL*). While the same trend is detected for inductions of gene expression by both technologies, in

the case of repressions, this is not always true (but compare section 5B). We found three elements whose sequences encoded senescence associated proteins (Table 11; *ests* 246D2T7, 246B12T7, and 212B17T7). They were more than 10-fold suppressed in *cim*205 based on microarray data.

Table 11: Comparison between gene expression quantification by the DNA microarray and by Northern (RNA) blot analysis.

Est/clone	*cim*205		*cim*328		*cim*713		Homology
	chip	Northern	chip	Northern	chip	Northern	
PR1	8.7[1]	18.9	8.3	3.0	6.1	12.7	
PR5	3.3	3.2	3.9	1.7	6.1	1.15	
PR2	3.8	2.4	2.8	1.1	5.2	2.0	
PDF1.2	1.4	0.8	1.4	0.3	1.6	0.5	
PAL	3.5	2.5	3.4	3.4	5.2	4.6	
NIM1	2.8	1.7	2.3	0.9	2.6	1.5	
CuZnSOD	1.4	1.5	2.1	1.8	3.3	1.0	
CAT2	1.2	1.0	1.1	2.1	1.4	1.5	
CAT3	1.3	3.3	1.2	2.8	1.5	2.7	
118N4T7	5.7	3.2	4.0	2.0	7.7	3.9	no homology
92I21T7	0.2	0.7	0.5	0.9	0.2	1.5	XET [4]
165P23T7	1.5	0.4	2.4	0.8	3.4	0.3	Brassinosteroid regul. Protein
129F12T7	3.5	1.5	3.5	1.0	5.5	1.5	squalene epoxidase
203C22T7	5.0	2.9	4.2	1.6	9.3	2.3	no homology
103L19T7	4.7	0.9	3.3	0.5	5.1	0.4	unknown
156F15T7	1.9	2.9	2.6	1.7	5.9	1.0	cp nucl. DNA bind. Protein [2]
246D2T7	0.1	2.0	0.6	0.9	0.9	3.0	Senesc ass prot [3]
246B12T7	0.1	0.5	0.5	1.1	1.1	2.5	Senesc ass prot [3]
212B17T7	0.1	3.0	0.5	0.9	1.0	-	Senesc ass prot [3]
E12G2T7	0.2	0.8	0.4	1.3	0.2	0.43	XET [4]
174N16T7	2.3	(1.2) [5]	2.6	(0.8)	5.9	(2.1)	LTP3
122B2T7	3.1	2.4	3.2	1.5	3.9	2.1	Chaperon 60β

1. For genes on the control plate, average of three values are indicated except for *CHS*, where two spottings failed
2. chloroplast nucleoid DNA binding protein
3. Senescence associated protein
4. Xyloglucan endotransglucosylase related protein
5. Northern blot quantifications were done with LTP1 as a probe, which is very similar to LTP3

This suppression was not confirmed to this extent by Northern blot analysis with any of the three *ests*. However, for *ests* 92I21T7 and E12G2T7 (both Xyloglucan endotransglucosidase related), the reduced gene expression was confirmed by both

technologies. Actually, in Northern blot analysis for this gene it was found that not only the quantity of this mRNA is reduced in *cim* mutants but also the length of the mRNA, possibly through alternative splicing. In general, it can be concluded that changes in gene expression less than 2-fold in microarray experiments are not significant without supplementary confirmation, but that stronger fold-inductions are in good accordance to data obtained by Northern blot analysis. The linear correlation coefficient between data obtained by Northern blot analysis and by DNA microarray, calculated from 40 data pairs, results in r = 0.84, which expresses a satisfactory linearity.

Analysis of the expression profiles of cim mutants

Besides one control plate and two Synteni control plates for normalization, 102 plates, each containg 96 *est* fragments, were spotted on the DNA microarray. Ideally, this could represent 9792 different genes. However, resequencing and BLAST comparisons revealed some redundancies in the *est* population, so that we estimate the actual number of monitored genes to about 5000. In cases were sequencing did not match the previously indicated identity of the clone, the name of the *est* was changed. In the hybridization experiments with the three *cim* mutants, 7815 (*cim*205), 8069 (*cim*713), and 7186 (*cim*328) elements gave readable signals. In six experiments listed in Table 9 (excluding experiment 4), 1562 different elements displayed at least in one hybridization a significantly altered (two-fold) change in expression, 1061 of those only in one experiment, 14 in all seven. 478 different elements displayed altered gene expression in at least one of the three *cim* mutants. As shown in table 12, in mutant *cim*713 many more genes than in the other mutants are differentially regulated (499 elements in *cim*713 versus 275 elements in *cim*205 and 180 elements in *cim*328 changed more than 2-fold). In NahG plants, the expression of 310 genes was changed and BTH application altered the expression of 916 genes more than 2-fold compared to untreated wild-type. Increasing the cut-off from a 2-fold difference to a 2.5 fold difference in gene expression reduces these numbers to less than half, and only few elements display a more than 3-fold change (table12). It should be noted, that for low-intensity read-outs (<1000 fluorescent signal intensity (FSI)) the correlation curve is biased towards cy5 and for results falling into these intensity regions, a 2-fold, or even 3-fold induction is not necessarily meaningful. We chose therefore a 2.5 fold cut-off for further analysis (see below). For interesting elements with low FSI, a careful analysis will be required. The distribution of the

FSI shows that most mRNAs fall into the class of low- to medium-abundant mRNAs (1 - 50 mRNA per cell; 1000 – 10,000 FSI). Only a few have intensities above 11,000 FSI, which corresponds to highly abundant transcripts (100 - 500 transcripts per cell, based on an estimated total number of 100,000 transcripts per cell; Kamalay and Goldberg, 1980). The largest changes in abundance were observed for low copy mRNAs that are usually more than 10-fold induced. The highly expressed housekeeping genes do not change the transcription rate in the *cim* mutants.

Table 12: Counts of elements displaying altered gene expression in three *cim* mutants, in plants treated with BTH and NahG plants. The total number of elements giving valid signals is shown, along with the number of elements with at least a 2-fold, 2.5-fold, 3-fold, or 4-fold difference in signal in the mutant (or treatment) compared to untreated wild-type. The five elements in *cim*328 that displayed a more than 4 fold change in expression are three elements for *PR-1*, one element for *PR-5* and *est* 203C22T7 (extensin).

Number genes induced	*cim*205	*cim*328	*cim*713	BTH	NahG
2-fold	275	180	499	916	310
2.5-fold	86	47	153	400	165
3-fold	45	19	79	261	120
4-fold	25	5	29	125	71
Total valid signals	7815	7186	8069	7816	7122

Because the majority of genes on the microarray stayed unchanged (under given experimental settings), we chose to limit the comparative analysis between the three mutants to genes that show altered transcription.

To compare the three mutants to each other, gene groups were created: Genes that changed significantly in one mutant (more than 2.5 fold) were grouped together and their gene expression data in the other mutants were obtained and plotted in profiles. The expression of 86 genes with differential expression in mutant *cim*205 was compared to their expression in *cim*713, *cim*328 and in the BTH control experiment. A characteristic of *cim*205 is the group of repressed genes around the *est* 246D2T7 and *est* 246B12T7 (both derived from genes encoding a senescence associated proteins). A third *est* derived from this gene family, *est* 212B17T7 is also strongly suppressed. This reduction in expression is not found in the other *cim* mutants, but is seen in the BTH treated tissue. As an example of a gene strongly induced

in *cim*205, *est* 203C22T7 (BLAST similarity: extensin) can be cited. In mutant *cim*205, as many genes are suppressed as induced.

This is not the case in mutant *cim*713. Most of the 153 selected genes with altered expression (cut-off 2.5-fold) in this mutant are upregulated and only a few are downregulated. It is evident that many more elements display significant alterations in gene expression in *cim*713 compared to *cim*205. Although the spectrum of gene inductions in the other two mutants looks very similar to the spectrum of *cim*713, those gene inductions are usually weaker and not necessarily significant. Several elements with the most dramatic changes in expression are annotated and the results of BLAST similarity searches are indicated in table 13. They include the known SAR genes, and genes encoding for cell-wall modifying proteins, such as extensins (*est* 118N4T7 and 203C22T7) and xyloglucan endotransglycosylase related proteins (*est* E12G2T7 and 92I21T7), as well as genes whose induction was not expected, such as the genes encoding squalene monooxygenase and a cytochrome P450, and also genes with unknown protein function (*est* 156F15T).

In mutant *cim*328, only 47 genes show more than 2.5 fold alterations in expression compared to wild-type. 35 genes are found only in the *cim*205 gene group (40%), and 107 genes are unique to the *cim*713 gene group (70%). In *cim*328, only 13 genes were found that did not also have altered expression in one of the other mutants. More than half of the genes induced in *cim*328 showed also altered expression in one or both of the other mutants: 11 elements induced in *cim*328 were also induced genes in mutant *cim*205, 6 are induced in mutant *cim*713 and 17 genes are common to all three gene groups. These 17 elements are listed in table 13.

Table 13: *ests* that revealed differential gene expression (fold expression over wild-type) in the *cim* mutants or following BTH treatment (2 days after 0.3 mM BTH treatment)

est	*cim*205	*cim*328	*cim*713	BTH	Blast similarity
203C22T7 [1]	5.0	4.2	9.3	6.3	extensin
212B17T7	-7.0	-2.0	1.0	-5.8	senescence associated protein
246D2T7	-7.9	-1.8	-1.1	-4.4	senescence associated protein
246B12T7	-7.9	-1.9	1.1	-5.5	senescence associated protein
118N4T7 [1]	5.7	4.0	7.7	8.4	no homology/extensin [3]
156F15T7	1.9	2.6	5.9	-	no homology
92I21T7	-4.2	-2.0	-6.1	-9.4	xyloglucan endotransglycosylase related
103L19T7 [1]	4.7	3.3	5.1	10.1	cystein protease inhibitor

174N16T7	2.3	-	6.3	-	non-spec. LTP
E12G2T7	-5.9	-2.6	-4.9	-11.1	xyloglucan endotransglycosylase related
122B2T7	3.1	3.2	3.9	5.7	chaperone 60β
129F12T7	3.5	3.5	5.5	9.2	squalene monooxygenase
PR1 (3x) [2]	10.8	10.0	9.0	13.9	(PR1)
PR2 (3x) [2]	4.7	3.4	6.2	8.5	(PR2)
PR5(3x) [2]	3.6	4.1	6.2	8.8	(PR3)
PAL [2]	3.5	3.4	5.2	7.1	(PAL)
CHS [2]	3.7	3.6	3.6	4.5	(CHS)
105D3T7	3.5	3.6	4.7	7.9	cytochrome P450

1 these genes fall in the class of the 17 genes with expressions altered more than 3-fold in all three mutants
2 highest value out of three is shown
3 est 118N4T7 is likely to be a chimeric clone

Interestingly, most of those common genes were already known as marker genes for disease resistance (e.g. *PR-1*), and all of them are also induced by BTH. Two notable exceptions are the previously mentioned genes for a cytochrome P450 (unknown substrate specificity) and for squalene monooxygenase. Mutant *cim*328 has also increased expression of extensin genes (*est* 118N4T7, 203C22T7) but only slightly reduced levels of senescence associated protein transcript.

Also, instead of selecting anonymous gene subgroups for comparison, one can also look at the expression of genes in specific pathways (Table 14). Expression of genes for proteins in the early steps of the Halliwell-Asada pathway had been studied in Northern blot analysis (see section 3.E.I.). On the DNA microarray, two more genes were added, ascorbate peroxidase (*APX*) and a glutathione peroxidase (*GLUPER*) gene. Both of them were consistently induced about 2-fold in the *cim* mutants over wild-type (table 10).

The microarray can thus be a powerful tool in dissecting pathways. In the analysis of mutants, gene expression profiles describe more precisely the characteristics than single marker genes do. The fingerprints for *cim* mutants show many similarities (the "SAR profile"), but also distinct features that might eventually help to explain phenotypic differences of the mutants.

Table 14: *est* coverage of genes encoding central enzymes of the phenylpropanoid pathway **(A)** and the Halliwell-Asada cycle **(B)**. *ests* that are present in the non-redundant set

of *ests* (used in part for the DNA microarray described above) are printed in bold, and the relative expression of the corresponding genes in the *cim* mutants compared to wild-type is indicated.

Gene	*est* clones (MSU collection)	Expression
(A) Phenylpropanoid pathway		
PAL1 and *PAL2* [1]	**CDNA, 128N22T7**, 187E19T7	3 – 5-fold induced
PAL3	**82B7T7, 154F7T7**, H2C1T7	No good signal
C4H	**240A11T7**, 123D6T7, 126E1T7, 118J19T7	Not spotted
4CL	**G2D3T7**, 175B13T7, 116E12T7, 69F10T7	Not spotted
Chalcone synthase	**CDNA, 177N23T7**, 187C23T7, 34H9T7, 43C1T7, 88D15T7	CDNA: 3 – 4-fold induced, *est*: no differential expression
Chalcone flavanone isomerase	**177A20T7, 240H10T7**	No differential expression
Dihydroflavonol-4-reductase	No *est*	n.a. [4]
Anthocyanidin synthase	No *est*	n.a.
Isoflavone reductase	**177A9T7, 122B24T7**	No differential expression
EPSPS [2]	**CDNA, 123B10T7**, G10B8T7, 62C8T7, 131D24T7	No differential expression
(B) Halliwell-Asada cycle		
Glutathione reductase	G2C3T7	Not spotted
Glutathione peroxidase	**CDNA** [3], **118F4T7**, 81B2T7, 218D7T7, 229J21T7, 118F6T7, 250D17T7, 99N6T7	2-fold induced
Dehydroascorbate reductase	No *est*	n.a.
Monodehydroascorbate reductase	**G7B8T7**, 38H7T7, 134D13T7	No good signal
Ascorbate peroxidase	**CDNA, 172F1T7, 186E1T7**, 109F7T7, 213G16T7, 90G17T7, 157B9T7, 251I20T7	1.8-fold induced
Superoxide dismutase Fe	E6E5T7	Not spotted
Cu/Zn	**CDNA** [3], 156O19T7	2-fold induced
Mn	**CDNA** [3], **200F22T7**, 123N9T7, 105G4T7, 109J19T7, 94E7T7	No differential expression

1. The sequences of *PAL1* and *PAL2* genes are too similar to identify gene-specific *ests*.
2. EPSPS is not part of the phenylpropanoid pathway, but of the shikimate pathway, furnishing the phenylpropanoid precursor phenylalanine.
3. cDNAs were not full-length clones, but only fragments as described in Table 4.
4. n.a.: not applicable

III. Two-dimensional protein gel electrophoresis of *cim* mutants

Bacteria possess, on average, four modified polypeptide chains for every three genes, in yeast, there are three polypeptides per gene and, in humans, between three and twenty different modified polypeptides can be found for each gene. Moreover, there is a poor correlation between the abundance of mRNA and their encoded proteins in eukaryotes, arguing for the crucial importance of both qualitative and quantitative analysis at the protein level (Landegren, 1999). In plant disease defense, post-transcriptional regulation on protein level has repeatedly been shown (Conrath *et al.*, 1997; Dietrich *et al.*, 1990). Two-dimensional (2D) polyacrylamide gel electrophoresis can resolve between 2500 and 10,000 proteins but the low loading capacity, unprecise pH gradients and difficult identification (either by comparison or by microsequencing) limit the actual use of this technology (Pennington *et al.*, 1997). 2D gel electrophoresis has been used before to characterize *Arabidopsis* mutants (Santoni *et al.*, 1994). We attempted to compare results obtained with the DNA microarray to changes in protein patterns. Total protein extracts of *cim*328, *cim*713, *cim*677, NahG, *cim*713xNahG, and BTH-treated wild-type plants were separated on 2D gel electrophoresis. The insufficient separation of proteins made it difficult to distinguish proteins in *cim* mutants that differ from wild-type. Several additional signals are marked that showed up in all three repetitions of the experiments. In total, about 10 signals that differ in each experiment compared to wild-type can be identified. In the picture showing the BTH control, one might expect to identify PR-1 (pI 8.5, 14.9 kDa). However, none of the signals were identified by microsequencing. It can be concluded that for subtle changes in expression, as found in *cim* mutants, 2D gel electrophoresis cannot provide enough sensitivity.

4. Mapping And Cloning Of *cim* Mutants

In order to clone the gene responsible for a Cim phenotype, a positional cloning approach was attempted. As a first step, the mutants had to be crossed to a genetically diverse ecotype and a linkage disequilibrium between the Cim phenotype and the genotype in a segregating F2 population had to be identified. For increased resolution, the number of F2 plants that must be scored increases. For many *cim* mutants, the expression of the Cim phenotype was unstable in mixed ecotype backgrounds. For four *cim* mutants, a rough map-

position (table 6) was obtained by screening through large F2 populations in order to isolate with confidence a sufficient number of F2 siblings with clear phenotype. Segregation in F2 progeny from the outcross of mutant *cim*713 most closely resembled a Mendelian segregation ratio; therefore this mutant was chosen for further cloning efforts.

A. Phenotyping of mutant *cim*713

There were a number of problems involved in the phenotyping of *cim*713. First, there is no visible, morphological phenotype associated with the *PR-1* overexpression phenotype. As an alternative, the increased disease resistance of *cim*713 was a potentially useful trait for identifying the Cim phenotype in the segregating F2 mapping population. However, there are no known pathogen isolates described that grow equally well on the parental *A. thaliana* ecotypes *Col-0* and *Ler* (these ecotypes were chosen because of the high rate of polymorphisms between the two genomes, estimated at 1 nucleotide polymorphism every 250 bp; Konieczny and Ausubel, 1993). An obvious trait to score in the F2 is the increased activity of *PR-1*/luciferase, the basis for the initial screen. There are several problems associated with this approach as well:

* The crossing partner *A. thaliana* ecotype *Ler* does not contain the *PR-1*/luciferase reporter gene, therefore one quarter of the F2 will not be usable for phenotyping. Plants containing the *PR-1*/luciferase transgene can be identified by selection on kanamycin or by PCR screening for the T-DNA. Although growth on GM plates under selection does not induce the *PR-1* gene expression, the phenotype changes slightly and phenotyping was never solely based on results obtained from those experiments. Therefore, a PCR for the luciferase gene was also established and run on all F2 plants and on at least 6 F3 progeny, to allow plants homozygous or heterozygous for the reporter gene to be distinguished.

* *PR-1* can be induced by various stresses. Although the levels of induction are usually small (with the exception of cell death-inducing events), the high sensitivity of the *PR-1*/luciferase system might make it difficult in some cases to distinguish between stress induction and genetic induction.

These problems can be partly circumvented by using endogenous biochemical markers, such as *PR-1* gene induction, or accumulation of salicylic acid, or the accumulation of PR-1

protein. These techniques were tried to find the most reliable technique to determine the phenotype of F2 plants. Because each F2 plant has to be confirmed in the next generation on population level (16 F3 plants are required to distinguish between heterozygosity and homozygosity of a dominant trait in F2 plants with 99% confidentiality; Koorneef and Stam, 1992) the technique must not be labor-intensive. Assuming proper segregation, the required number of F2 plants to map a gene in the *Arabidopsis* genome to a 50 kb region, can be derived from the corresponding average resolution of 0.25 cM. With the size of the *Arabidopsis* genome round off to 100,000 kb and 500 cM, on average, 5 cross-over (c.o.) per haplotype per meiosis occur. The required population size to find 1 cross-over per 50 kb results with 100.000 kb/50 kb = 2000 cross-overs in the population to 2000 c.o. / 5 c.o. per haplotype = 400 haplotypes = 200 plants in the F2 population. For the probability to find one recombinant in this interval to be no less than P = 95%, with E = 200

$$n = \ln(1-P)/\ln(1-1/E)$$
$$= \ln(1-0.95)/\ln(1-1/200)$$
$$= 600 \text{ F2 plants}$$

600 F2 plants would have to be screened and confirmed in the F3 on population level.

This number excludes the determination of SA accumulation by organic extractions. Western blot analysis (or ELISA) for the PR-1 protein was tried with the strongest mutant *cim*810. The detection is less sensitive than for the *PR-1* mRNA. Furthermore, it also requires an effort per sample too big for the vast sample number required.

The detection of the *PR-1* marker gene expression by dot blot (RNA) analysis is feasible in the F2, and was used to confirm in random samples the phenotype determination by luciferase activity but does not help to improve the reliability of the phenotype determination. The rate of phenotype miscalls was the same as when F2 plants were scored for luciferase activity. The Cim phenotype is only expressed in a small percentage of the F2 plants, the penetrance is incomplete (15% vs. 75% expected).

Hence, only the luciferase *in vivo* assay would allow the screening of 10,000 F3 plants in a reasonable amount of time. Because of the penetrance problem encountered in the *Ler* ecotype, about 5000 F2 plants were screened for constitutive luciferase expression. 855 F2 plants, for which the phenotype could be called without doubt, were used for mapping purposes. Mostly, they were homozygous for the *cim* mutation. For the 855 F2 plants, the

presence of the luciferase gene was confirmed by PCR. The F3 populations were scored on both kanamycin selection and without selection.

To avoid the severe penetrance problem of the mutation in the *Ler* ecotype, mapping crosses to the *A. thaliana* ecotype *Wassilewskija* (*Ws-0*) were also tried, but the penetrance did not improve.

B. Genetic mapping of *cim*713

Several genetic markers have been identified, including SSLP and CAPS markers, that are polymorphic between the *A. thaliana* ecotypes *Col-0* and *Ler*. To get a rough map-position, about 30 markers were used on a F2 population of *cim*713 of 65 individuals. Linkage to the phenotype was found with the SSLP marker nga280 (at 81.4 cM on the Lister and Dean RI map, 2 recombinants in 124 meiosis). The next closest SSLP markers, nga248 (at 40.0 cM) and nga111 (at 111.4 cM) showed less linkage. In this genomic region, no PCR-based genetic marker was available. RFLP markers were present in the interval, but these are difficult to use on individual F2 *Arabidopsis* plants because of the amount of genomic DNA required. To convert RFLP markers into PCR-based markers, several RFLP probe clones were sequenced and sequence-specific primers were designed. PCR fragments from both parental ecotypes were digested using 48 to 80 different restriction endonucleases to find a cleaved amplified polymorphic sequence (CAPS). Polymorphisms were detected as differences of fragment sizes after electrophoretic separation. This procedure was successful for the RFLP markers mi209, mi304 and mi291a, and for the gene *NIA2*. No polymorphism was detected for RFLP marker mi106. When the F2 population was scored at the marker mi291a, a tight linkage to the phenotype (5 recombinants in 120 meiosis) was detected. Using the Kosambi function (map distance D in cM: $D = 25 \ln (100 + 2r)/(100 - 2r)$, where r is the recombination frequency expressed as a percentage (Koorneef and Stam, 1992; Kosambi, 1944); for recombination frequencies r smaller than 4%, the difference between map distances D and r becomes less than 1% and r can be used without logarithmic corrections) to convert recombination frequencies into map-distances, the distance between the two genetic markers mi291a and nga280 was calculated at 9.5 cM ($r = 5.7$), which corresponds well to the published value (11 cM). PAP240 was identified as an expressed sequence tag hybridizing to a YAC clone in this region. The clone (obtained from M. Raynal, INRA, Perpignan, France)

was converted into a CAPS marker in similar manner to the RFLP markers. The marker PAP240 divided the genetic distance between the markers mi291a and mi209, and narrowed the interval containing the *cim* mutation to roughly 2 cM.

Sequencing the clones of the two RFLP markers APK100 and LOX in this interval did not detect any polymorphism between *Ler* and *Col-0* DNA. The clones were therefore only used as RFLP probes on DNA blots of pooled F3 progeny from individual F2 plants showing recombination between PAP240 and mi209. The two markers delimited an interval of 6 and 4 recombinants from the *cim*713 mutation. This corresponds to a map-distance of 6 / (2 x 855) = 0.35 cM on the left side (APK100) and 0.23 cM on the right side (LOX). Ideally, an interval of 0.58 cM corresponds to a physical distance of 100 kb in *Arabidopsis*, and can be spanned by one or two BAC clones.

C. Physical mapping

The *Arabidopsis* genome is well represented in several large-capacity vectors. Three YAC and two BAC libraries exist that are partly assembled into contigs. From the physical map, sequence information can be derived to design new genetic markers and eventually to construct a high-resolution genetic map. Sequence information can be obtained from publicly available BAC end sequences (http://genome.bio.upenn.edu), YAC end rescue, cloning of BAC (random or end-) fragments, or by the identification of *ests* that hybridize to BAC clones. With the rapid progress of the *Arabidopsis* sequencing project, the chance of finding sequence information of entire BAC clones is also increasing.

The two closest flanking genetic markers, APK100 and LOX were used as anchor points on the physical map. They were hybridized to BAC filters containing subsets of the two available BAC clone libraries, IGF and TAMU (ABRC stock center). APK100 hybridized to BAC F16J8, F15I10, F22G10 and F8H4. The LOX probe hybridized to BAC T7N22, T3A10, and to the BAC clones F19C6, F9I9, F26H12 and F5P9. From these starting points, a physical contig was constructed, using both experimental and non-experimental data. J. Ecker (Univ. Pennsylvania, PA), and T. Altmann's laboratory (Max-Planck-Institut für molekulare Pflanzenphysiologie, Golm, Germany) provide hybridization data of BAC clones to BAC end fragments and construct BAC contigs (http://www.mpimp-golm.mpg.de/101/mpi_mp_map/bac.html). The considered region was however not yet

contiged. Washington University (St. Louis, MO) provides *Hin*dIII fingerprints of BAC clones (http://genome.wustl.edu/gsc/cgi-bin/arab/atdatabase.shtml). Based on similarity in restriction fragment patterns, different BAC clones can be aligned with certain probabilities. This allows a virtual walk on the chromosome and enabled us to integrate the two genetic markers into one physical contig. The contig was confirmed by PCR of BAC end fragments on overlapping BAC clones. The minimal pass between the two markers requires most likely three BACs (F2H24, T28H10, and F14C21, however, the overlap between T28H10 and F14C21 is uncertain. The physical distance between the two genetic markers is therefore unlikely to exceed 350 kb (average size of a BAC clone is 100 kb).

D. Genetic marker development and fine-mapping

The sequence of one of the BAC clones in the contig, BAC F20D21, became partly available (Genbank accession number AC005287; preliminary release from October 3rd 1998). The first attempts to develop CAPS markers from PCR fragments were unsuccessful, revealing the presence of a non-polymorphic genomic region. Therefore, the fragments were used as probes in RFLP survey blots. Probe 20D21-13 detected a *Dra*I restriction polymorphism between the two parental lines *Col-0* and *Ler*. The 10 recombinants between the two flanking markers were scored at this locus and a cosegregation of this marker with the mutation *cim*713 was found (table 15). Based on the theoretical resolution obtainable with 1710 meiosis that were used to build the linkage map and assuming an even disribution of recombination, the mutation should lie within 23 kb of any cosegregating marker.

Because the development of new markers in this region proved to be extremely difficult by the restriction digest approach, alternative methods were also tried. Simple sequence repeats, like polydT, or polydCA, were identified on the BAC F20D21 and flanking primers were designed to identify length polymorphisms (SSLP). Usually, these repeats are meiotically unstable and vary in size between evolutionarily distinct ecotypes.

Six of these short fragments were amplified from both ecotypes and separated on high resolution gels, but no length polymorphisms were detected.

As a last, most accurate (and costly) solution, systematic sequencing of genomic DNA from both parents corresponding to the insert cloned in BAC F20D21 was conducted. PCR primers were designed every 1 kb, the slightly overlapping fragments purified and sequenced

with the same PCR primer pair. Of the 100 kb of BAC F20D21 that were sequenced, only 7 single nucleotide polymorphisms between the ecotypes *Col-0* and *Ler* were detected, 5 of them were restriction fragment polymorphisms and 4 were converted into CAPS markers (orf5, 20D21-2, orf52, cf2-12). The fifth polymorphism was already used as the RFLP marker 20D21-13. This is an extremely low rate of genetic polymorphism and explains the difficulties encountered in marker development by random trials.

Interestingly, all these markers cosegregated with the *cim*713 phenotype, thus extending the interval with no recombination event in the mapping population to at least 140 kb, which is more than 5 times larger than we estimated for our F2 population size based on genome-average recombination.

In order to mark the limits of the interval, markers that reveal recombinations between the mutations and the nearest defined genetic marker loci were needed. The BAC clone T22H22, which overlaps with the right (SP6) end of the BAC F20D21 was, similar to F20D21, partly sequenced by the *Arabidopsis* genome initiative in the course of this work (AC005388). PCR fragments on BAC T22H22 were generated of both parental lines and sequenced. The locus T26 (at 26 kb from the BAC end F20D21) was polymorphic and the recombinant F2 plant number 1006 had a recombination event between this marker and the mutation, thus limiting the physical interval on the right side (table 15).

To limit the interval on the left side of BAC F20D21, all available BAC end sequences between APK100 and F20D21-2 were used in Southern blot analysis to find restriction length polymorphisms (the BAC end sequences were too short to develop CAPS markers), but no such polymorphism was detected. Therefore other mismatch detection methods, used primarily in mammalian mapping projects and routine identification of known mutations, were tried.

Heteroduplex analysis detects changes in confirmation of DNA duplexes caused by single base pair mismatches. PCR fragments in the corresponding genomic region, are denatured, mixed and hybridized with PCR products of the second parental ecotype. Any point mutation results in the formation of two heteroduplexes as well as two homoduplexes. The heteroduplexes have altered confirmations, which can be detected by the altered migration in a polyacrylamide gel. Hauser *et al.* found in *Arabidopsis* 50% of 36 loci (230 bp to 1000 bp PCR fragments) to be polymorphic (Hauser *et al.*, 1998). Although we reproduced

the published polymorphisms, we did not identify new markers by this method using short PCR fragments that were derived from the available BAC end sequences in the interval of interest. Similarly, we did not detect poolymorphisms with a second method, the RNase cleavage assay that is supposed to detect 90% of all mismatches (Ambion, Austin, TX). Specific PCR primers for 5 BAC end fragments between the markers APK100 and 20D21-2 with T7 and SP6 promoter sequences were used for thermocycle amplification, followed by an *in vitro* transcription and a cross hybridization of the transcripts of the two ecotypes. Different endoRNases were used in vain to cleave potential mismatches. Therefore, as before, sequencing seemed to be the last solution to identify the rare nucleotide polymorphisms in this genomic region between the two ecotypes *Col-0* and *Ler*. Using long (26-mer) PCR primers from several BAC end sequences in the direction of BAC F20D21, a long range PCR fragment between BAC end F17M20 and the left (T7) end of BAC F20D21 was amplified and cloned into the vector pCR2.1. These additional 9.1 kb were sequenced by random transposon integration. Marker development based on this sequence is in progress that will hopefully limit the cosegregating interval on the left side of BAC F20D21.

The bad genetic resolution left us with a region of around 150 kb, in which the *cim* gene should be localized. In order to identify possible candidate genes, open reading frames (ORF) on BAC F20D21 and T22H22 were identified using consensus prediction programs (Genscan, http://ccr.081.mit.edu/Genscan.html). BAC F20D21 contained 28 putative genes. For some, the prediction was confirmed by the presence of *ests* in the databank or by Northern blot analysis. ORFs were sequenced in the *Col-0* wild-type and *cim*713 mutant. In cases where the sequenced fragments did not overlap with the next fragment, Northern blot analysis was performed to detect potential point mutations in promoter elements, leading to changes in gene expression. One gene with homology to a TMV resistance protein showed a 5-fold induction of expression in the mutant *cim*713. The expression of a serine/threonine protein kinase in this interval was 3-fold reduced in the mutant.

Table 15: Recombinant F2 plants between the genetic markers PAP240 and nga280 on chromosome 1 in the mapping population of *cim*713.

recomb	PAP 240	APK 100	20D2 1-2	20D2 1-13	orf5	pheno	orf52	cd2-12	T26	LOX	nga 128	PLR B910	mi 209	nga 280
242	H	H	H	H	H	H	H	H	H	C	C	C	C	C
250	H	C	C	C	C	C	C	C	C	C	C	C	C	C
253	H	C	C	C	C	C	C	C	C	C	C	C	C	C
255	H	H	n.d.[2]	L	L	L	L	L	L	L	L	L	L	L
265	L	H	H	H	H	H	H	H	H	H	H	H	H	H
296	H	L	L	L	L	L	L	L	L	L	L	L	L	L
382[1]	C	C	C	C	n.d.	C	n.d.	C	C	n.d.	H	H	n.d.	H
507	H	C	C	C	C	C	C	C	C	C	C	C	C	C
513	H	C	C	C	C	C	C	C	C	C	C	C	C	C
524	C	C	C	C	C	C	C	C	C	C-H	H	H	H	H
536	H	C	C	C	C	C	C	C	C	C	C	C	C	C
550	C	C	H	H	H	H	H	H	H	H	H	H	H	H
563	H	H	C	C	C	C	C	C	C	C	C	n.d.	n.d.	C
573	H	C	C	C	C	C	C	C	C	C	C	C	C	C
582	C	C	C	C	C	C	C	C	C	C	H	C	H	H
666	H	C	C	C	C	C	C	C	C	C	C	C	C	C
705	C	C	C	C	C	C/H	C	C	C	C	nd	H	H	H
711	H	C	C	C	C	C	C	C	C	C	C	C	C	C
743	H	C	C	C	C	C	C	C	C	C	C	C	C	C
837	C	C	C	C	C	C	C	C	C	C	C	C	H	H
870	H	H	C	C	C	C	C	C	C	C	C	C	C	C
929	H	H	C	C	C	C	C	C	C	C	C	C	C	C
1006	C	C	C	C	C	C	C	C	H	H	H	H	H	H
1036	C	C	C	C	C	C	C	C	C	H	H	H	H	H
1051	H	C	C	C	C	C	C	C	C	C	C	C	C	C
1063	H	C	C	C	C	C	C	C	C	C	C	C	C	C
1096	H	n.d.	C	C	C	C	C	C	C	C	C	C	C	C
1136	C	L	L	L	L	L	L	L	L	L	L	L	L	L
1137	C	C	H	H	H	H	H	H	H	H	H	H	H	H
1169	C	C	C	C	C	C	C	C	C	C/H	H	H	H	H
1175	H	C	C	C	C	C	C	C	C	C	C	C	C	C
1193	C	C	C	C	C	C/H	n.d.	H	H	C/H	H	H	H	H

1 recombinant 382 did not set seeds in the F2, therefore no data on RFLP markers were obtained.
2 n.d.: not determined

Other genes did not show alterations in expression. In the two genes with altered expression, no mutations on nucleotide level were detected in the coding region or in the promoter. The changes might therefore be a consequence rather than the cause of the mutant phenotype. Although almost the entire cosegregating region on BAC F20D21 and T22H22 was sequenced, no (point) mutation was identified in mutant *cim*713.

Two additional mRNA species were detected using the long range PCR fragment on BAC F17M20 as a probe in Northern blot analysis, and neither was differentially expressed in the mutant. The presence of two genes in this interval was also predicted by the Genscan program after the sequence had been obtained. One of the genes was identified as a homolog of the mammalian ataxin gene. The long range PCR fragment has not yet been sequenced for the mutant *cim*713.

E. Cloning of flanking DNA in *cim* mutants generated by T-DNA tagging

A second screen for *cim* mutantswas performed in a population of mutants generated by T-DNA tagging and seven additional *cim* mutants were identified (table 16). Segregation analysis for Basta resistance in segregating T2 populations revealed that for most of the mutants, several T-DNAs were integrated into the genome. For lines 27001, 28492, and 33673, no progeny plant died in populations comprising at least 50 individuals in the selection for Basta resistance, which suggests that at least three T-DNA inserts may be segregating independently.

The number of inserts was also estimated by probing genomic Southern blots from each of the mutants with T-DNA probes (using pBluescript and the RB). This method usually gave a higher estimate of T-DNA inserts than segregation analysis. There are two possible reasons for this discrepancy. Multiple T-DNAs may have inserted at a single genetic locus, or partial T-DNAs, lacking the *BAR* gene but containing pBluescript and RB segments may have been inserted.

Table 16: Genetic analysis of *cim* mutants generated by T-DNA insertion. Out of 10,000 primary transformants, 7 lines were retained that showed reproducible *PR-1*/luciferase activity in the next generations. Those were submitted to segregation analysis on selective media (Basta) and Southern blot analysis to identify T-DNAs that cosegregate with the phenotype. Plasmid rescue, or TAIL PCR was performed to clone flanking genomic DNA.

line	luciferase activity [1]	number inserts (segregation)	luciferase segregation in T2 (+:-)	number inserts (Southern)	TAIL (T)/ plasmid rescue (P)
27001	+++	>3 [2]	1:3	4 - 5	P, T
33553	+++	2	1:3	1	n.t.[3]
28492	++	>3 [2]	1:3	2 - 3	T
32053	+++	>2	1:3	3	T
25949	++	1-2	1:7	3	n.t.
30881	++	2	1:2	3	T
33673	++	>3 [2]	1:3	3 - 4	T

1. Luciferase *in vivo* activity was compared to mutant *cim713* (set to ++). Stronger activity was rated +++.
2. No plant died (out of more than 50 T2 plants)
3. n.t.: not tried

In many cases, no specific T-DNA was correlated to the expression of the Cim phenotype. Because the mutations are recessive or codominant, the plants without (or with a weak) Cim phenotype might be heterozygous at the *cim* locus, thus giving a signal in the Southern blot analysis that does not match a visible phenotype. Alternatively, T-DNA-induced, but not linked, genomic rearrangements might cause the observed phenotype.

For line 27001, plasmid rescue was tried for the RB flanking sequence with an *Eco*RI, or *Hin*dIII digest. The left border can not be rescued because of the lack of unique restriction sites.

Several cloned flanking genomic fragments were sequenced and the inserts compared by BLAST similarity search to sequences in the *Arabidopsis* genome or elsewhere. Two flanking sequences were derived from sequenced parts of the *A. thaliana* genome, two had homology to *Arabidopsis ests* and two had no significant homology. One plasmid contained only the T-DNA cloning vector, probably integrated into the genome by inefficient cleavage of the T-DNA at the border sequence during the transformation.

For five lines, a TAIL PCR was performed to get LB- and RB-flanking genomic DNA. For all lines, TAIL PCR was successful and several fragments per line were either subcloned

or directly submitted for sequencing. Some fragments contained only the T-DNA sequence, possibly because of tail-to-head or head-to-head cointegrations of several T-DNAs into the genome. In several cases, known genes or sequenced genomic regions were cloned. The fragments were used in Southern blot analysis to confirm the insertions and to identify the gene knock-out that cosegregated with the phenotype. These genes will be transformed into the mutants to establish wild-type phenotypes and to genetically prove the cloning of *cim* genes.

5. Epistasis studies with *cim* mutants

Several other mutants involved in disease resistance became available during the course of this study. Most of them were loss-of-function mutations and for understanding the genetic signaling pathway it is important to establish the epistatic relationship between these mutants and the *cim* mutants. However, several problems occur:

- The Cim phenotype is not expressed in 100% of the progeny, making genetic analysis in F2 populations difficult, especially if no genetic marker linked to the mutation can be followed.
- Several loss-of-function mutants are not in the same ecotype. For instance the *eds1* mutant exists in the *A. thaliana* ecotype *Ler* and *Ws-0*. This renders analysis of resistance more difficult because of the presence of race-specific *R* genes (no *Peronospora parasitica* isolate has been identified that infects both *Col-0* and *Ler*). In addition, crosses to other ecotypes increase the penetrance problem, as described previously.

Despite these problems, crosses of several *cim* mutants to the disease resistant compromised mutants *nim1-4*, *pad4*, *eds1* and *ndr1*, as well as to the jasmonic acid signaling-deficient mutant *jar1* and the ethylene-insensitive *etr1* mutant have been obtained. They can be analyzed after a *cim* gene has been cloned so that the genotype can be confirmed in the F2. Crosses to the SA-degrading NahG line could be analyzed in the F1, because both, the NahG phenotype and the Cim phenotype are inherited in a (co-) dominant manner.

A. Analysis of *cim* mutants in NahG backgrounds

All *cim* mutants displayed increases in SA content, with the exception of the "weakest" mutant *cim*658. To address whether SA accumulation is required for the primary Cim phenotype (i.e. *PR* gene expression), all 16 *cim*s were crossed to the SA-degrading NahG line and monitored for *PR-1*/luciferase gene expression in the F1 by *in vivo* photon emission. F1 plants were confirmed for double marker antibiotic resistance by plating on selective media (kanamycin for presence of the luciferase marker gene, and hygromycin for the presence of the *NahG* gene) and by RNA blot analysis for the hygromycin gene. *PR-1*/luciferase gene expression was suppressed in all crosses (at least three crosses per mutant with 10 to 20 plants per cross), revealing the dependence of *PR-1* gene expression on prior SA accumulation.

Resistance to *P. parasitica Noco2* was, surprisingly, not completely abolished in at least two mutants (*cim*695 and *cim*713) in combination with NahG (mutants *cim*8, *cim*658, and *cim*677 were not tested). It is unlikely that this resistance is due to residual SA because *PR-1* gene expression was suppressed and because other mutants that accumulated 2-fold more SA than mutant *cim*713 nevertheless lost the *Noco*2 resistance phenotype in a NahG background (e.g. mutant *cim*810). Gene expression profiles for F1 plants before and after pathogen treatment showed that constitutive transcription of the endogenous *PR-1* gene is indeed suppressed. Similarly, the *NIM1/NPR1* gene, which is usually activated two- to threefold during an SAR response, was not induced in presence of NahG. These results demonstrate that SA accumulation is not a general requirement for *Peronospora* resistance in at least two *cim* mutants. Interestingly, we did not observe a correlation of SA-independent resistance and *PDF1.2* gene expression, as previously suggested (Bowling *et al.*, 1997). Mutant *cim*205 expressed *PDF1.2* to higher levels than wild-type, *cim*824 did not. Both mutants are susceptible to *P. parasitica* in a NahG cross. Mutant *cim*695 does not constitutively express the *PDF1.2* gene, yet *PDF1.2* transcription was induced when this mutant was crossed to a NahG line. In contrast, mutant *cim*713 expressed elevated *PDF1.2* mRNA levels that were decreased in a NahG background, but resistance to *P. parasitica* was retained in both mutants. Thus, transcription of neither SAR genes nor *PDF1.2* (as a putative marker for ethylene/jasmonate mediated inducible resistance) is increased in *cim*713 when crossed to NahG, revealing a novel, unrelated mechanism for resistance to *P. parasitica*.

B. Gene profiling for *cim* mutants in NahG backgrounds

In order to correlate the SA-independent disease resistance mechanism of mutant *cim*713 to marker genes, and to monitor gene expression changes of *cim* mutants in a NahG background, three DNA microarray experiments were conducted (experiments 5, 6, and 7 in table 9). Using the results of these experiments, the gene expression of the mutants *cim*205 and *cim*713 in a NahG background can be compared to the expression profile of both the NahG plants, and to profiles of the mutants themselves (using information of the experiments 1 and 3 in table 9). 280 elements that showed altered gene expression (2.5 fold over wild-type) in at least one of the experiments were chosen for comparison. Compared to previous profiles, the strong differences in gene expression are remarkable. Several genes were strongly upregulated in NahG (up to 40-fold), others show up to 100-fold reductions in gene expression in NahG compared to wild-type. In general, these genes are also altered in the *cim* mutants in NahG backgrounds (although NahG is not a mutant *sensu stricto*, the crosses of *cim* mutants to the NahG line are in the following chapter for the ease of wording referred to as "double mutants".) Interestingly, most of those genes have no homology to entries in public databases (table 17). As with the *cim* mutants (table 11), data obtained by Northern blot analysis correspond well with microarray data, which are usually somewhat higher (table 17). The overall pattern of gene induction in NahG compared to the double mutants revealed that NahG imposes many changes in gene expression such that the three profiles are all very similar. Again, as with the *cim* mutants alone, a common subset of genes with similar expression patterns in the double mutants and NahG (the "NahG profile") can be identified, including the elements with the most extreme changes, *est* 90G10T7 (+41.4), 140E3T7 (+23.1), 198G3T7 (-26.3), 127P22T7 (-67.3) and 103C7T7 (-21.5). There are nevertheless differences between the profiles. For instance, *cim*205 x NahG expresses more than 30 genes 5- to 10-fold stronger than wild-type that are not differentially expressed in NahG or *cim*713 x NahG. About 20 genes that are reduced between 5- to 10-fold in NahG are either not as strongly repressed in the double mutants or are not repressed at all. Thus, there are differences between the two double mutants and NahG.

Table 17: *ests* that revealed differential gene expression in *cim205* x NahG and *cim713* x NahG on the DNA microarray ("chip"), and the values obtained in Northern blot analysis ("blot").

est/clone	homology	BTH		NahG		*cim205*xNahG		*cim713*xNahG	
		blot	chip	blot	chip	blot	chip	blot	chip
198G3T7	none	11.3	5.6	-8.8	-26.3	-16.5	-40.6	-19.5	-40.1
103C7T7	unknown	3.7	2.3	-10.7	-21.5	-6.0	-13.8	-12.1	-28.8
CXC750 [1]		5.5	3.4	-1.7	-10.5	-9.1	-65.3	-10.6	-19.1
93C1T7	unknown	5.1	3.0	-1.5	-97.2	-6.2	-22.0	-8.6	-13.4
124K13T7	CXC750	5.1	3.4	-1.7	-65.3	-7.6	-10.5	-12.5	-19.1
127P22T7 [2]	unknown	2.3	5.1	-1.2	-67.5	-1.3	-66.3	-1.3	-30.0
140E3T7	unknown	1.6	-1.2	1.4	23.1	5.4	17.9	28.8	23.6
119D12T7	none	4.3	-1.6	2.1	21.7	4.1	9.5	3.6	17.8
119C13T7	none	1.7	-	1.3	9.9	3.6	10.8	2.1	12.8
90G10T7	none	1.8	1.1	1.3	41.4	4.3	29.5	2.4	38.0
103L19T7	unknown	15.4	10.1	-2.0	-3.9	-2.9	-4.7	-2.5	-2.8
PR2		6.5	n.t. [3]	-6.1	n.t.	-13.5	n.t.	-8.3	n.t.
PR5		7.9	n.t.	-4.1	n.t.	-10.5	n.t.	-4.7	n.t.
PAL		2.4	n.t.	-3.7	n.t.	-9.6	n.t.	-5.6	n.t.
172A5T7	none	5.9	n.t.	-9.0	n.t.	-10.4	n.t.	-13.2	n.t.
163B24T7	none	4.6	n.t.	-16.8	n.t.	-13.9	n.t.	-26.8	n.t.
105D3T7	unknown	7.9	n.t.	-8.6	n.t.	-3.0	n.t.	-11.6	n.t.
178N5T7	none	4.3	n.t.	-6.7	n.t.	-10.7	n.t.	-8.7	n.t.
177C23T7	none	5.8	n.t.	-7.2	n.t.	-5.2	n.t.	-10.1	n.t.

1. not the same *ests* were used in Northern blot analysis and DNA microarray

2. Northern blot signal very weak

3. n.t.: not tested

The second question addresses the similarity between gene expression in the mutants compared to the mutants in NahG backgrounds. It is already clear that fundamental changes in gene expression are introduced with the *NahG* gene. Many genes that are induced in *cim* mutants (including *PR-2* and *PR-5*) were repressed by the presence of NahG. One of the strongest changes in expression occurred for the gene *CXC750* that is pathogen-inducible (Aufsatz and Grimm, 1994) and whose expression is obviously SA-dependent (table 17). The function of CXC750 is still unknown. *PR-1* gene expression is not repressed by NahG compared to wild-type because its basal levels in wild-type are already low. Others (most prominently *est* 140E3T7, 198G3T7, 103C7T7) did not change the expression in *cim* mutants but are unique to the NahG line and the double mutants. And a few genes retained the same

expression pattern in both the *cim* mutant and the double mutants (for example some elements around the *est* 93C1T7). Of the 200 genes that are altered more than 2.5-fold in the cross *cim*205 x NahG, 56 are also altered more than 2.5 fold in *cim*205 alone, but in only one case is gene expression changed in the same direction. All the other genes were altered inversely. This complementary effect on gene expression of the *cim* mutants and the NahG lines might be directly compared by eliminating the intrinsic base of comparison (wild-type). Subtracting the expression values of one experiment from another (Δ[*cim* x NahG; wt] - Δ[*cim*; wt] = Δ[*cim* x NahG; *cim*]) shows that only very few differential expression values are close to zero, i. e. most of the expression values of *cims* and NahG lines are different. Clearly, the changes in gene expression imposed by NahG are radical and do not only concern genes required for catechol degradation. Salicylic acid might have a major importance for the regulation of many genes. Again, in the 2D protein gel electrophoresis these changes are not well reflected.

6. Using Expression Profiling To Study The Transcriptome of *Arabidopsis* During SAR

To gain a more thorough understanding of the complex regulation of gene expression during SAR, and to identify new (marker) genes associated with SAR, we used expression profiling (Schena *et al.*, 1995; Shalon *et al.*, 1996). PolyA+ RNAs obtained from plants grown under 18 different conditions that either induce or repress SAR (see Table 18 below) were used in mixed hybridizations against RNA of untreated wild-type plants on a DNA microarray with 10,000 ESTs, a Unigene set representing roughly one third of all *Arabidopsis genes.*

Hybridizations were done by Synteni, Inc., Fremont, CA as described by Ruan, *et al.*, 1998. For spiking controls, refer to http://www.synteni.com/client/controls/index.html. We included several quality controls to check the accuracy of the DNA microarray data. Spot-to-spot variability was estimated using 30 cDNAs spotted three times each. Measured variability was usually smaller than 20%. The inverse labeling of sample *cim11* not only showed the reproducibility of independent experiments (linear regression coefficient r = 0.83), but also the symmetry of labeling. We also compared more than 40 DNA microarray data points to Northern blot results and found a linear correlation of r = 0.83.

The EST "Unigene" set was obtained from the *Arabidopsis* Biological Resource Center, (Ohio State University, Columbus, OH). 10,000 of the 14,000 clones were amplified by PCR using the M13 universal reverse and -21 forward primer, with modified 5' amino end for spotting onto a glass slide. A fraction of the ESTs, including all mentioned in this manuscript, were resequenced at our facility. Roughly 30 % of the EST sequences did not correspond to the original sequence as represented in the AATDB. Based on our own BLAST search and on estimations made by others (Delseny *et al.*, 1997), we extrapolated the redundancy of the "Unigene" set to 1.5 to 2-fold. Hence, while the majority of the checked genes match only one EST in our set, many match at least two ESTs. One explicit example can be found in the PR1 regulon: Disregarding the three CDNA clones PR1, PrxC, and PR5 that were added purposely several times on the microarray, 19 ESTs represent 14 different genes.

In addition, two treatments which alter plant metabolism, but are not related to SAR, were analyzed (Table 18). Some samples were taken during the induction phase of SAR (4 hours after BTH induction), or from primary, pathogen-infected tissue, but most of the samples were derived from (steady state) SAR maintenance phase (e.g. 48 hours after BTH induction, cim mutants). Together, the experiments gave us 1.8×10^5 gene expression data points. Under these 18 conditions, however, only 660 different genes displayed significant differential expression compared to wild-type in at least two SAR-relevant samples (significance level defined as 2.5 fold differences from wild-type). By comparing the fluorescence signal intensities to spiked controls, abundance of mRNA species can be determined . Most of the genes with altered expression fall into the class of low-abundance transcripts (1 to 10 copies per cell), while a few were in the medium abundance class. Housekeeping genes, with more than 100 copies per cell, did not exhibit altered gene expression under SAR.

Table 18: Diversity of conditions used to describe the transcriptome of *Arabidopsis thaliana* during SAR.

Sample (all compared to wild-type tissue)	Comments
Adenylosuccinate synthetase antisense Hydantocidin (0.9 mM)	Reference inductions, irrelevant to resistance
cim6 *cim7* *cim11*	Constitutive SAR mutants
NIM1 overexpresser	Primed SAR response
NahG *cim6NahG* *cim11*NahG NahG + *Pst* DC3000 *avrrpm1* 2' tissue, 44 hours	No SA accumulation, no SAR response
NahG + *Pst* DC3000 *avrrpm1* 1' tissue, 44 hours BTH 4 hours BTH 48 hours	Primary (partial) LAR induction SAR induction SAR maintenance
P.p. EMWA1, 48 hours *Pst* DC3000 *avrrpm1* 1' tissue, 44 hours *Pst* DC3000 *avrrpm1* 2' tissue, 44 hours *nim* + Pst DC3000 *avrrpm1* 2' tissue, 44 hours *P.p.* Noco5, 48 hours	Compatible (disease causing) interaction Incompatible interaction

First, we compared gene induction patterns of these 660 genes among each of the different conditions. Data representing a series of differential gene expression measurements was obtained. Analysis of both the regulation of individual genes under varying conditions or of the overall similarity of the various conditions based on expression profiles required the calculation of distance matrices. These were calculated in S-Plus using the dist function under the euclidean metric. Clustering of similarly regulated genes was performed and visualized in S-Plus using hierarchical clustering under the compact method with the, functions hclust and plclust. Trees used to depict similarity in gene expression patterns among the conditions may be produced using the fitch and drawtree programs distributed in the Phylip suite (Phylogeny Inference Package) version 3.57c.

Out of 10,000 ESTS, 661 that displayed differential gene expression under two or more of the 16 SAR-inducing or -repressing conditions were used to calculate similarities in gene expression profiles. To indicate a scale of similarity between two different conditions, we introduced a unit of similarity, which corresponds to the summed difference of relative gene inductions under two conditions over all 660 genes, calculated according to the formula: $Sum_{genes1\text{-}n}(Abs(\Delta(expression_{cond1}, expression_{cond2})))/n$ with n = number genes (661). This simplified scale compares well to the results obtained by the more complex statistical methods

and resulted for the pair *(cim7, cim11)* in an average difference of gene induction of 1.1 units per gene, compared to 3.7 units for the condition pair *(cim11,* NahG).

Genetic, chemical and biological induction of SAR all caused very similar patterns of gene induction. The three *cim* mutants, which constitutively express SAR, have a pattern of transcriptional induction similar to that caused by BTH during SAR maintenance. Interestingly, the overexpression of *NIM,* which primes the SAR response, also resembles SAR maintenance (U.S. Patent No. 6,031,153; Cao *et al.*, 1998; Maleck *et al.*, 1998). Compatible and incompatible races of the pathogens *Peronospora parasitica and Pseudomonas syringae* induced many common genes, although the pattern of gene induction was not identical to either of the BTH-treated samples. In contrast to the requirement for SA accumulation to high levels for SAR induction, many gene inductions are triggered by low levels of SA. For example, in NahG-expressing plants that accumulate only low levels of SA, the expression of more than 300 genes is significantly altered. This emphasizes a general regulatory role for SA in the plant cell.

NahG suppresses SAR gene expression in crosses to two of the SAR-constitutive *cim* mutants, *cim6 and cim11,* to a baseline resembling that of NahG-expressing plants. In addition, NahG expression results in a characteristic gene expression fingerprint in secondary tissue from plants inoculated in primary tissue with avirulent bacteria. This corresponds to the inability of these plants to establish SAR. Interestingly, the corresponding primary tissues in NahG-expressing plants display changes in gene expression which compares very closely to wild-type primary, infected tissue and this sample does not cluster with other NahG samples. This is consistent with findings from grafting experiments that SA is required for SAR development in systemic tissue, but that the systemic signal can be emitted from primary infected NahG-expressing tissue (Delaney *et al.*, 1994; Gaffney *et al.*, 1993; Vernooij *et al.*, 1994).

We used these expression profiles to identify classes of genes that were consistently coregulated, and we sought to identify genes associated with SAR. We used statistical phylogenetic cluster analysis to compare the gene expression profiles among 660 genes under all 18 conditions. Clusters of genes with similar induction patterns under all conditions were grouped. The two clusters most distinct from the PR1 gene induction pattern, with EST 127P22T7 and EST 163B24T7 as "type genes," contain genes that are strongly (more than 20

fold) suppressed by NahG expression. The cluster containing chalcone synthase (ChS) as the type gene includes genes induced during SAR-maintenance, but not during SAR-induction; these are weakly repressed by NahG expression. The cluster containing EST 209E19T7 defines genes that are transcriptionally induced in NahG-expressing plants. The cluster containing EST 118P18T7 defines genes that are not significantly responsive to SAR-inducing conditions like chemical and genetic induction, but do respond to avirulent bacteria and are downregulated in NahG expressing plants. Phenylalanine ammonia lyase (PAL) and 20 other ESTs that cluster together are repressed by NahG expression, but are induced during the maintenance phase of SAR, for example in *cim* mutants or 48 hours after BTH treatment. The cluster of "PR1 like" genes exhibits similar induction behavior to genes in, the PAL gene cluster but these genes are only weakly suppressed in NahG-expressing plants.

To address the question of expression induction relatedness among genes within one cluster, we analyzed the PR1 cluster in detail. The PR1 regulon contained 25 other ESTs (17 different genes). These are prime candidates for SAR marker genes and the encoded proteins are likely to play a physiological role in SAR. The estimated 1.5 to 2-fold redundancy of our EST set is a good internal control for this analysis and we also included three replicates of the PR5 and the PerC cDNAs (as well as 28 other relevant cDNAs) on the DNA-microarray. All three copies of the two genes cluster with PR1, showing the robustness of the DNA microarray analysis. Similarly, two ESTs that are derived from Asparagine synthetase, two from the gene for blue copper binding protein and two from glutathionine S-transferase (GST) cluster with PR1. The standard deviations for differences in expression of the 30 members of the PR1 cluster were smaller than one unit of differential gene expression (compared to 2.5 units cut-off for our definition of significant alterations) under all 18 conditions. Thus, genes in one cluster show highly reproducible expression in our analysis.

To help explain the molecular basis of this co-regulation, we analyzed the promoters of those eight PR1 cluster ESTs encoded within sequenced genomic regions. Strikingly, these genes all shared a common TTGAC consensus cis-element known to bind proteins of the WRKY zinc finger transcription factor binding family, often in multiple copies. In control promoters of eight ESTs selected randomly among the 634 non-PR1-like ESTs, the element TTGAC was only found in the statistically expected frequency. The WRKY element was not present in three of the randomly selected promoters, in three promoters once, in one promoter

twice, and in one promoter three times. WRKY elements are sufficient for defense gene transcriptional induction in different plant systems (Lebel *et al.*, 1998). A functional analysis of the BTH-induction of the PR1 promoter revealed two elements sharing all or four out of five consensus WRKY sites at -676 and - 638 (Lebel *et al.*, 1998). The former acts as a repressor and the latter as a SA responsive element. It is therefore likely that WRKY transcription factors are responsible for the common induction pattern of the eight genes listed below in Table 19, and possibly also for the remaining 22 ESTs of the PR1 gene cluster for which no genomic sequence is presently available.

We did not identify another common consensus sequence in the eight PR1 cluster promoters analyzed. The NFκB binding site found in the PR1 promoter which is required for full BTH-induction is not conserved among the eight promoters within the PR1 regulon.

Table 19: Genes in the PR1-like cluster contain a common cis-element. For the 26 ESTs clustering with PR1, genomic sequence was obtained, if available. 1 kb fragments upstream of the EST end or the predicted gene (if annotated) were scanned for common promoter elements. Numbering of the upstream sequences starts at the 5' end of the longest available EST, except for the PR1 promoter, for which primer extension has been performed (Lebel *et al.*, 1998).

EST	homology	cis element	position	Genomic sequence of EST or gene
cDNA	PR1	ATTGAC TATGAC ATTGAC CTTGAC ATTGAC	-676 to -671 -638 to -643 -544 to -539 -492 to -487 -181 to -176	AF096294 (PR1 promoter)
134C20T7	No homology	ATTGAC TTTGAC TTTGAC	-799 to -794 -735 to -731 -335 to -330	ABO17065, 16134 - 15793
192K7T7	No homology	TTTGAC	-841 to -836	AC002330, 99468 - 99711
147P23T7	Myrosinase binding protein	TTTGAC	-857 to-852	AC001645, 7088 - 7298
245M18T7	Tyrosine aminotransferase	TTTTGAC CTTTGAC	-653 to -648 -109 to -104	AC006585, 81840 - 83922
121I3T7	No homology	GTTGAC TTTGAC	-996 to -991 -792 to -787	ABO11485, 48839 - 48492
179C14T7	7-ethoxy-coumarin-O-deethylase	ATTGAC TTTGAC CTTGAC TTTGAC	-892 to -887 -813 to -808 -745 to -740 -18 to -13	AC003680, 33897 - 33546

184F12T7	Late embryonic abundant protein (SAG21)	TTTGAC CTTGAC TTTGAC TTTGAC	-847 to -842 -671 to -666 -117 to -112 -36 to -31	AF069298, 63830 - 64550
WRKY consensus		TTGAC		Dorey *et al.*, 1998

In addition to identification of common regulatory elements, cluster analysis of expression profiles provides a tool to derive physiological functions of genes. This is important for sequences with no close homologs in the databank (for example EST 134C2OT7 or EST 192 K7T7) and also for genes with structural similarity to genes with known function (such as asparagine synthetase). We anchored the results of our statistical analysis first to those genes that were known to be co-regulated with PR1. Interestingly, several genes in the PR1 cluster encode proteins involved in redox regulation. Although reactive oxygen intermediates do not likely act as secondary messengers in SAR downstream of SA accumulation, they are clearly implicated in defense response, either as local second messenger molecules, as direct cellular and microbial toxins or in cell-wall cross-linking and other oxidative processes (Alvarez *et al.*, 1998). From the several known *Arabidopsis* peroxidases, peroxidase C seems to be specifically induced under plant defense conditions, as are some GSTs. Supporting this is the previous demonstration that both peroxidase C and GST are transcriptionally activated in a superoxide-dependent manner in the *Arabidopsis* mutant *Isd1*, which lacks the ability to halt hypersensitive-response-like lesions (Jabs *et al.*, 1996; Dietrich *et al.*, 1994). Likewise, the blue-copper protein might be involved in the regulation of redox stages during SAR or might be the chelator of free cellular copper used for Cu/Zn SOD. Cu/Zn SOD is suppressed during SAR (Fodor *et al.*, 1997).

It has been proposed that plant metabolism adapts to pathogen attack (Batz *et al.*, 1998). We can now more precisely identify such changes and investigate their importance in disease resistance. For example, it is instructive to speculate about the possible role of the coregulation of glutamine-dependent asparagine synthetase with the PR1 cluster during SAR. An increased flux of carbon through the phenylpropanoid pathway (lower overall N:C ratio) might liberate nitrogen, thus creating a demand for asparagine synthetase as a detoxifying enzyme. Similarly, pathogen infection increases the level of free ammonia in the plant, following an increased metabolism of proteins (Batz *et al.*, 1998). Asparagine synthetase can

also utilize ammonia as a substrate. Hence, our observation that asparagine synthetase is co-regulated with PR1 focuses attention on physiological changes during SAR that might not be directly linked to defense. Alternatively, the induction of asparagine synthetase might be related to the induction of tyrosine transaminase. This is also the first time that the production of toxic mustard oils, released by myrosinase from glucosinolates, can be correlated to SAR in *Arabidopsis.*

Our analysis describes the first map of the plant defense transcriptome during SAR in *Arabidopsis.* Using cluster analysis of DNA microarray data, significant changes in gene induction can be differentiated from random correlations. The large number of data points filters out coincidental co-induction and enhances the resolution and significance of serial analysis of expression profiles. It is clear that the description of complex signaling networks can only be obtained by investigating multiple inductive conditions, and not, as before, by pair-wise comparisons. Thus, similar analysis of additional mutants and inducing conditions will further enlarge the complexity and refine the resolution of the entire *Arabidopsis* transcriptome.

7. Using Expression Profiling To Identify Inducible Promoters

By comparing gene expression changes across various treatments, groups of co-regulated genes (regulons) were identified and the genomic sequences of genes within the regulons were examined to identify common sequence motifs likely to act as regulatory elements. These regulatory elements may then be used to make promoters that drive controlled gene expression. As described above, gene expression profiling using DNA microarrays was used to identify groups of genes that show a similar pattern of expression in response to biotic and abiotic stimuli, especially the 18 biotic and abiotic inducers of SAR described above in Table 18. Expression profile data uncovered groups of genes responsive to exogenous factors but not endogenous signals.

Using PCR select and microarray gene chip technology described above, genes were identified that are responsive to BTH and/or pathogens. Furthermore, we compared expression profiles in response to biotic and abiotic inducers of SAR and assessed the requirement for salicylic acid and the *NIM1* gene for mRNA accumulation. By analysis of data

generated with cDNA microarrays, sets of genes that are responsive specifically to exogenous application of BTH were identified (see Table 20 below). The discovery of a set of BTH-inducible genes that are not responsive to SA or pathogens was especially surprising given the expectation from previous studies of the SAR signal transduction pathway that BTH would always act as a functional analog of SA and would also activate the SAR response in the same manner as pathogen infection. The regulatory regions from these differentially expressed genes can be isolated using conventional cloning techniques and used as specifically inducible promoters, such as BTH-specific promoters.

Table 20: Induced genes based on >5x BTH induction in a wild-type *Arabidopsis* line at 4, 24, and 48 hours post-BTH and in the *nim1-4* mutant *Arabidopsis* line at 4 and 24 hours post-BTH. These are candidate genes for BTH-specific promoters that are induced by BTH but not SA or pathogens. Genes below the double line are induced by BTH in the *nim1-4* mutant, but are not induced by SA or pathogen. Gene names (EST IDs) in **bold** are induced more than 5x by BTH and are also induced in the *nim1-4* mutant line by BTH (NIM1 independent genes). x: not in cluster analysis; # indicates cluster analysis result. Numbers indicate fold-induction relative to control treatment. If no number is present, fold-induction is less than 2.5.

Gene Name (EST ID)	Cluster	Sequence verified	Wildtype BTH 4h	Wildtype BTH 24h	Wildtype BTH 48h	*nim1-4* BTH 4h	*nim1-4* BTH 24h
H2D10T7	x			6.5	5.6		
91B11T7	x			14.9	5.9		
90N23T7	x					6.6	
241N24T7	x		7.7			5.7	
240N6T7	x		5.3				
229F6T7	x		7.5			8.4	
223M15T7	x		5.2			7.3	
194G1T7	3	Y				6.5	
190M3T7	x		5.4				
190C24T7	3	ACO02333		5.1			
177C23T7	4	P48980		5.2	5.8		
175H5T7	9	Y		5.8			
168A5T7	x				5.9		
167I19T7	8	Y		5.2			
166P4T7	8	Y		6.1			
166N10T7	4	Y			5.1		
163L17T7	x			5.2	6.1		

162J11T7	0	Y	7.4				
159O2T7	8	Z49697		5.1			
159O16T7	8	Y		6.5			
158A15T7	x		5.3				
141D8T7	x				5.3		
134G3T7	x				6.1		
134F18T7	x				8		
126N11T7	10	Y			7.1		
125H21T7	8	Y		6.7			
124A5T7	4	103D24T7			6.4		
122F9T7	x			6.4	8.5		
122B7T7	8	Y			6.8		
122B24T7	4	Y			5.6		
122B23T7	8	ACOo4484		5.7	6.3		
122A22T7	8	121J12T7		6.9			
121H22T7	x			10.2			
121G2T7	7	Y			5.8		
121G21T7	8	Y		5.1			
121B21T7	8	Y			5.6		
119N23T7	x			9			
118L22T7	x			12.4	6.6		
118K13T7	x			8.3			
117P7T7	x			11.3			
117P14T7	8	118L7T7		8.6	7.8		
117P13T7	8	Y		9.8	5.2		
117O8T7	x				5.8		
117N21T7	4	S66345			5.9		
117G12T7	8	Y		5.2			
117E1T7	x			9	5.3		
117E18T7	9	118K3T7		6.3			
117D3T7	x			9.4			
117D21T7	x			9.2	5.4		
111F20T7	0	Y			5.7		
90N23T7	x		3.7	2.3	1.2	6.6	3.9
241N24T7	x		7.7	1.1	-1.6	5.7	2.8
229F6T7	x		7.5	2.4	1.1	8.4	4.5
228G7T7	3	Y	4	2.4	3.2	2.2	3.6
223M15T7	x		5.2	2.3	1.3	7.3	3.3
222N23T7	x		4	2.4	2.2	2.1	3.8
222C9T7	7	Y	3.6	-3.1	1.4	1.1	14
202O18T7	x		3.1	-1	-1.2	4.3	1.6
194G1T7	3	Y	3.7	-1.2	-2.6	6.5	1.4
191H15T7	x		3.1	2	1.6	1.3	3.2
177H9T7	x		3.9	2.3	1.9	2.1	3.8
172I13T7	x		3.2	2	-1.4	1.9	5
162J3T7	7	Y	3.8	-1.9	2.1	1.5	4.8
162J11T7 (N -gene)	0	Y	7.4	-1.1	5	1.2	3.7

155E20T7	x		3	3.1	2.3	1.1	4.2
125K7T7	x		4.8	1.2	-1.4	3.1	1.6
109K9T7	x		-7.8	2.3	3.5	-10.3	5
107A4T7	x		3.9	-1.3	-2.3	3.7	1.7

REFERENCES

The references cited herein are indicative of the current state of the art. Each of the following is incorporated by reference into the instant disclosure.

U.S. Patent No. 6,057,490

U.S. Patent No. 6,031,153

U.S. Patent No. 5,614,395

International Patent Application WO 98/03536

Aarts, N., Metz, M., Holub, E., Staskawicz, B. J., and Daniels, M. J. (1998). Different requirements for *EDS1* and *NDR1* by disease resistance genes define at least two *R* gene-mediated signaling pathways in *Arabidopsis*. *Proc. Natl. Acad. Sci. USA* **95**, 10306-10311.

Adam, L., and Somerville, S. C. (1996). Genetic characterization of five powdery mildew disease resistance loci in *Arabidopsis thaliana*. *Plant J.* **9**, 341 - 356.

Agrios, G. N. (1988). *Plant pathology*, 3rd Edition (San Diego, CA: Academic Press).

Akama, K., Puchta, H., and Hohn, B. (1995). Efficient *Agrobacterium*-mediated transformation of *Arabidopsis thaliana* using the *bar* gene as selectable marker. *Plant Cell Rep.* **14**, 450 - 454.

Albertsen, H. M., Abderrahim, H., Cann, H. M., Dausset, J., Le Paslier, D., and Cohen, D. (1990). Construction and characterization of a yeast artificial chromosome library containing seven haploid human genome equivalents. *Proc. Natl. Acad. Sci. USA* **87**, 4256 - 4260.

Alexander, D., Goodman, R. M., Gut-Rella, M., Glascock, C., Weymann, K., Friedrich, L., Maddox, D., Ahl-Goy, P., Luntz, T., Ward, E., and Ryals, J. (1993). Increased

tolerance to two Oomycete pathogens in transgenic tobacco expressing pathogenesis-related protein 1a. *Proc. Natl. Acad. Sci. USA* **90**, 7327-7331.

Alexander, D., Stinson, J., Pear, J., Glascock, C., Ward, E., Goodman, R. M., and Ryals, J. (1992). A new multigene family inducible by tobacco mosaic virus or salicylic acid in tobacco. *Mol. Plant Microbe Interact.* **5**, 513-515.

Alonso-Blanco, A., Peeters, A., Koornneef, M., Lister, C., Dean, C., van den Bosch, N., Pot, J., and Kuiper, M. (1998). Development of an AFLP based linkage map of *Ler*, *Col*, and *Cvi Arabidopsis thaliana* ecotypes and construction of a *Ler/Cvi* recombinant inbred line population. *Plant J.* **14**, 259-271.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol.Biol.* **215**, 403 - 410.

Alvarez, M. E., Pennell, R. I., Meijer, P., Ishikawa, A., Dixson, R., and Lamb, C. (1998). Reactive oxygen intermediates mediate a systemic signal network in the establishment of plant immunity. *Cell* **92**, 773 - 784.

Asada, K. (1994). Production and action of active oxygen in photosythetic tissue. In Causes of photooxidative stress and amelioration of defense systems in plants, C. H. Foyer and P. M. Mullineaux, eds. (Boca Raton: CRC Press).

Aufsatz, W., and Grimm, C. (1994). A new, pathogen-inducible gene of *Arabidopsis* is expressed in an ecotype-specific manner. *Plant Mol. Biol.* **25**, 229 - 239.

Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., and Struhl, K. (1987). *Current protocols in molecular biology* (New York: J. Wiley and Sons).

Bachem, C. W. B., van der Hoeven, R. S., de Bruijn, S. M., Vreugdenhill, D., Zabeau, M., and Visser, R. G. F. (1996). Visualization of differential gene expression using a novel method of RNA fingerprinting based on AFLP: Analysis of gene expression during potato tuber development. *Plant J.* **9**, 745 - 753.

Baeuerle, P. A., and Baltimore, D. (1988). Activation of DNA-binding activity in an apparently cytoplasmic precursor of the NF-kB transcription factor. *Cell* **53**, 211 - 217.

Bate, N. J., Orr, J., Ni, W., Meromi, A., Nadler-Hassar, T., Doerner, P. W., Dixon, R. A., Lamb, C. J., and Elkind, Y. (1994). Quantitative relationship between phenylalanine ammonia-lyase levels and phenylpropanoid accumulation in transgenic

tobacco identifies a rate-determining step in natural product synthesis. *Proc. Nat. Acad. Sci. USA* **91**, 7608 - 7612.

Batz, O., Logemann, E., Reinold, S., and Hahlbrock, K. (1998). Extensive reprogramming of primary and secondary metabolism by fungal elicitor or infection in parsely cells. *Biol. Chem.* **379**, 1127 - 1135.

Bechtold, N., Ellis, J., and Pelletier, G. (1993). *In planta Agrobacterium* mediated gene transfer by infiltration of adult *Arabidopsis thaliana* plants. *Life Sciences* **316**, 1194-1199.

Bell, C. J., and Ecker, J. R. (1994). Assignement of 30 microsatellite loci to the linkage map of *Arabidopsis*. *Genomics* **19**, 137 - 144.

Belvin, M. P., and Anderson, K. V. (1996). A conserved signaling pathway: The *Drosophila Toll-Dorsal* pathway. *Annu. Rev. Cell Dev. Biol.* **12**, 393 - 416.

Bent, A. F., Innes, R. W., Ecker, J. R., and Staskawicz, B. J. (1992). Disease development in ethylene-insensitive *Arabidopsis thaliana* infected with virulent and avirulent *Pseudomonas* and *Xanthomonas* pathogens. *Mol. Plant-Microbe Interact.* **5**, 372-378.

Bent, A. F., Kunkel, B. N., Dahlbeck, D., Brown, K. L., Schmidt, R., Giraudat, J., Leung, J., and Staskawicz, B. J. (1994). *RPS2* of *Arabidopsis thaliana*: a leucine-rich repeat class of plant disease resistance genes. *Science* **265**, 1856-1860.

Berger, S., Bell, E., Sadka, A., and Mullet, J. E. (1995). *Arabidopsis thaliana Atvsp* is homologous to soybean *VspA* and *VspB*, genes encoding vegetative storage protein acid phosphatases, and is regulated similarly by methyl jasmonate, wounding, sugars, light and phosphate. *Plant Mol. Biol.* **27**, 933 - 942.

Bevan, M., and The European *Arabidopsis* genome project (1998). Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of *Arabidopsis thaliana*. *Nature* **391**, 485 - 488.

Bhalla, U. S., and Iyengar, R. (1999). Emergent properties of networks of biological signaling pathways. *Science* **283**, 381 - 387.

Birnboim, H. C., and Doly, J. (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucl. Acid. Res.* **7**, 1513 - 1523.

Blackstock, W. P., and Weir, M. P. (1999). Proteomics: quantitative and physical mapping of cellular proteins. *Trends Biotech.* **17**, 121 - 127.

Bohlmann, H. (1994). The role od thionins in plant protection. *Crit. Rev. Plant Sci.* **13**, 1-16.

Boller, T. (1991). Ethylene in Pathogenesis and Disease Resistance. In The Plant Hormone Ethylene, A. K. Mattoo and J. C. Suttle, eds. (Boca Raton, FL: CRC Press), pp. 293-314.

Botella, M. A., Coleman, M. J., Hughes, D. E., Nishumura, M. T., Jones, J. D. G., and Somerville, S. C. (1997). Map positions of 47 *Arabidopsis* sequences with sequence similarity to disease resistance genes. *Plant J.* **12**, 1197-1221.

Bouchez, D., and Hofte, H. (1998). Functional genomics in plants. *Plant Physiol.* **118**, 725 - 732.

Bowler, C., van Montagu, M., and Inze, D. (1992). Superoxide dismutase and stress tolerance. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **43**, 83 - 116.

Bowling, S. A., Clarke, J. D., Liu, Y., Klessig, D. F., and Dong, X. (1997). The *cpr5* mutant of *Arabidopsis* expresses both *NPR1*-dependent and *NPR1*-independent resistance. *Plant Cell* **9**, 1573 - 1584.

Bowling, S. A., Guo, A., Cao, H., Gordon, A. S., Klessig, D. F., and Dong, X. (1994). A mutation in *Arabidopsis* that leads to constitutive expression of systemic acquired resistance. *Plant Cell* **6**, 1845-1857.

Briggs, S. P. (1995). Grand unification theory in sight. *Curr. Opin. Biotechnol.* **5**, 128 - 131.

Britt, A. B. (1999). Molecular genetics of DNA repair. *Trends Pl. Sci.* **4**, 20 - 25.

Broekaert, W. F., Terras, F. R. G., Cammue, B. P. A., and Osborn, R. (1995). Plant defensins: novel antimicrobial peptides as components of the host defense system. *Plant Physiol.* **108**, 1353 - 1358.

Bullock, W. O., Fernandez, J. M., and Short, J. M. (1987). XL1-Blue: a high efficiency plasmid transforming RecA *Escherichia coli* strain with beta galactosidase selection. *BioTechniques* **5**, 376 - 379.

Burke, D. T., Carle, G. F., and Olson, M. V. (1987). Cloning of large segments of exogenous DNA into yeast by means of artificial chromosome vectors. *Science* **236**, 806 - 812.

Büschges, R., Hollrichter, K., Panstruga, R., Simons, G., Wolter, M., Frijters, A., van Daelen, R., van der Lee, T., Diergaarde, P., Groenendijk, J., Topsch, S., Vos, P., Salamini, F., and Schulze-Lefert, P. (1997). The barley *MLO* gene: A novel control element of plant pathogen resistance. *Cell* **88**, 695 - 705.

Cameron, R. K., Dixon, R., and Lamb, C. (1994). Biologically induced systemic acquired resistance in *Arabidopsis thaliana*. *Plant J.* **5**, 715-725.

Cao, H., Bowling, S. A., Gordon, S., and Dong, X. (1994). Characterization of an *Arabidopsis* mutant that is nonresponsive to inducers of systemic acquired resistance. *Plant Cell* **6**, 1583-1592.

Cao, H., Glazebrook, J., Clarke, J. D., Volko, S., and Dong, X. (1997). The *Arabidopsis NPR1* gene that controls systemic acquired resistance encodes a novel protein containing ankyrin repeats. *Cell* **88**, 57 - 63.

Cao, H., Li, L., and Dong, X. (1998). Generation of broad-spectrum disease resistance by overexpression of an essential regulatory gene in systemic acquired resistance. *Proc. Natl. Acad. Sci. USA* **95**, 6531-6536.

Cardon, G. H., Frey, M., Saedler, H., and Gierl, A. (1993). Mobility of the maize transposable element *En/Spm* in *Arabodopsis thaliana*. *Plant J.* **3**, 773 - 784.

The *C. elegans* Sequencing Consortium. (1998). Genome sequence of the nematode *C. elegans*: a platform for investigating biology. *Science* **282**, 2012 - 2018.

Century, K. S., Shapiro, A. D., Repetti, P. P., Dahlbeck, D., Holub, E., and Staskawicz, B., J. (1997). *NDR1*, a pathogen-induced component required for *Arabidopsis* disease resistance. *Science* **278**, 1963 - 1965.

Century, K. S., Holub, E. B., and Staskawicz, B. J. (1995). *NDR1*, a locus of *Arabidopsis thaliana* that is required for disease to both a bacterial and a fungal pathogen. *Proc. Natl. Acad. Sci. USA* **92**, 6597 - 6601.

Chapple, C. C. S., Shirley, B. W., Zook, M., Hammerschmidt, R., and Somerville, S. C. (1994). Secondary metabolites in *Arabidopsis*. In *Arabidopsis*, E. Meyerowitz and C. Somerville, eds.: Cold Spring Harbor Laboratory Press), pp. 989 - 1030.

Cho, *et al.* (1998a). Parallel analysis of genetic selections using whole genome oligonuclotide arrays. *Proc. Natl. Acad. Sci. USA* 95: 3752-3757.

Cho, *et al.* (1998b). A genome-wide transcriptional analysis of the mitotic cell cycle. *Molecular Cell* 2: 65-73.

Choi, S., Creelman, R. A., Mullet, J. E., and Wing, R. (1995). Construction and characterization of a bacterial artificial chromosome library of *Arabidopsis thaliana*. *Plant Mol. Biol. Rep.* **13**, 124 - 128.

Church, G. M., and Gilbert, W. (1984). Genomic sequencing. *Proc. Natl. Acad. Sci. USA* **81**, 1991-1995.

Clarke, J., Liu, Y., Klessig, D., and Dong, X. (1998). Uncoupling *PR* gene expression from *NPR1* and bacterial resistance: Characterization of the dominant *Arabidopsis cpr6-1* mutant. *Plant Cell* **10**, 557-569.

Conrath, U., Silva, H., and Klessig, D. F. (1997). Protein dephosphorylation mediates salicylic acid-induced expression of PR-1 genes in tobacco. *Plant J.* **11**, 747-757.

Creelman, R. A., and Mullet, J. E. (1997). Biosynthesis and action of jasmonates in plants. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **48**, 355 - 381.

Crute, I., Beynan, J., Dangl, J., Holub, E., Mauch-Mani, B., Slusarenko, A., Staskawicz, B., and Ausubel, F. (1994). Microbial pathogenesis of *Arabidopsis*. In *Arabidopsis*, E. M. Meyerowitz and C. R. Somerville, eds.: Cold Spring Harbor Laboratory Press), pp. 705.

Dagert, M., and Ehrlich, S. D. (1979). Prolonged incubation in calcium chloride improves the competence of *Escherichia coli* cells. *Gene* **6**, 23 - 28.

Dangl, J. L., Dietrich, R. A., and Richberg, M. H. (1996). Death don't have no mercy: Programmed cell death in plant-microbe interactions. *Plant Cell* **8**.

Davis, K. R., and Hammerschmidt, R. (1993). *Arabidopsis thaliana* as a model system for plant-pathogen interactions (St. Paul, Minn.: American Phytopathological Society).

Debener, T., Lehnackers, H., Arnold, M., and Dangl, J. L. (1991). Identification and molecular mapping of a single *Arabidopsis thaliana* locus determining resistance to a phytopathogenic *Pseudomonas syringae* isolate. *Plant J.* **1**, 289-302.

Dekker, J. (1995). Development of resistance to modern fungicides and strategies for its avoidance. In Modern selective fungicides, H. Lyr, ed. (Jena: Fischer), pp. 23 - 38.

Delaney, T., Friedrich, L., and Ryals, J. (1995). *Arabidopsis* signal transduction mutant defective in chemically and biologically induced disease resistance. *Proc Natl. Acad. Sci., USA.* **92**, 6602-6606.

Delaney, T., Uknes, S., Vernooij, B., Friedrich, L., Weymann, K., Negrotto, D., Gaffney, T., Gut-Rella, M., Kessmann, H., Ward, E., and Ryals, J. (1994). A central role of salicylic acid in plant disease resistance. *Science* **266**, 1247-1250.

Dellaporta, S. L., Wood, J., and Hicks, J. B. (1983). A plant DNA minipreparation. Version II. *Plant Mol. Biol. Rep.* **1**, 19 - 21.

Delseny, M., Cooke, R., Raynal, M., and Grellet, F. (1997). The *Arabidopsis thaliana* cDNA sequencing projects. *FEBS letters* **405**, 129 - 132.

DeRisi, J. L., Iyer, V. R., and Brown, P. O. (1997). Exploring the metabolic and genetic control of gene expression on a genomic scale. *Science* **278**, 680 - 686.

Desprez, T., Amselem, J., Caboche, M., and Hoefte, H. (1998). Differential gene expression in *Arabidopsis* monitored using cDNA arrays. *Plant J.* **14**, 643 - 652.

Diatchenko, L., Lau, Y. C., Campell, A. P., Chenchik, A., Moqadam, F., Huang, B., Lukyanov, S., Lukyanov, K., Gurskaya, N., Sverdlov, E. D., and Siebert, P. D. (1996). Suppression subtractive hybridization: a method for generating differentially regulated or tissue-specific cDNA probes and libraries. *Proc. Natl. Acad. Sci. USA* **93**, 6025 - 6030.

Dietrich, A., Mayer, J. E., and Hahlbrock, K. (1990). Fungal elicitor triggers rapid transient and specific protein phosphorylation in parsley cell suspension cultures. *J. Biol. Chem.* **265**, 6360-6368.

Dietrich, R. A., Richberg, M. H., Schmidt, R., Dean, C., and Dangl, J. L. (1997). A novel zinc finger protein Is encoded by the *Arabidopsis LSD*1 gene and functions as a negative regulator of plant Cell death. *Cell* **88**, 685 - 694.

Dietrich, R. A., Delaney, T. P., Uknes, S. J., Ward, E. R., Ryals, J. A., and Dangl, J. L. (1994). Arabidposis mutants simulating disease resistance response. *Cell* **77**, 565-577.

Dixon, R. A., Harrison, M. J., and Lamb, C. J. (1994). Early events in the activation of plant defense responses. *Ann. Rev. Phytopathol.* **32**, 479 - 501.

Dixon, R. A., and Lamb, C. J. (1990). Molecular communication in interactions between plants and microbial pathogens. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **41**, 339-367.

Dixon, R. A., and Paiva, N. L. (1995). Stress-Induced Phenylpropanoid Metabolism. *Plant Cell* **7**, 1085 - 1097.

Dong, X., Mindrinos, M., Davis, K. R., and Ausubel, F. M. (1991). Induction of *Arabidopsis* defense genes by virulent and avirulent *Pseudomonas syringae* strains and by a cloned avirulence gene. *Plant Cell* **3**, 61-72.

Dorey, S., *et al.* (1998). *Mol. Plant Microbe Interact.* **11**, 1102.

Duckett, J. G., and Read, D. J. (1991). The use of the fluorescent dye, 3,3'-dihexyloxacarbocyanine iodide, for selective staining of ascomycete fungi associated with liverwort rhizoids and ericoid mycorrhizal roots. *New Phytol.* **118**, 259 - 272.

Ebel, J., and Cosio, E. G. (1994). Elicitors of plant defense responses. *Internatl. Review of Cytology* **148**, 1 - 36.

Ecker, J. R. (1998). Genes blossom from a weed. *Science* **391**, 438.

Eisen, *et al.* (1998). Cluster analysis and display of genome-wide expression patterns. *Proc. Natl. Acad. Sci. USA* 95: 14863-14868.

Enyedi, A. J., and Raskin, I. (1993). Induction of UDP-glucose: salicylic acid glucosyltransferase activity in tobacco mosaic virus-inoculated tobacco (*Nicotiana tabacum*) leaves. *Plant Physiol.* **101**, 1375-1380.

Enyedi, A. J., Yalpani, N., Silverman, P., and Raskin, I. (1992). Localization, conjugation and function of salicylic acid in tobacco during the hypersensitive reaction to tobacco mosaic virus. *Proc. Natl. Accd. Sci. USA* **89**, 2480-2484.

Epple, P., Apel, K., and Bohlmann, H. (1997). ESTs reveal a multigene family for plant defensins in *Arabidopsis thaliana. FEBS Letters* **400**, 168-172.

Epple, P., Apel, K., and Bohlmann, H. (1997). Overexpression of an endogenous thionin enhances resistance of *Arabidopsis* against *Fusarium oxysporum*. *Plant Cell* 9, 509 - 520.

Epple, P., Vignutelli, A., Apel, K., and Bohlmann, H. (1998). Differential induction of the *Arabidopsis thaliana Thi2.1* gene by *Fusarium oxysporum* f. Sp. *matthiolae. Mol. Plant Microbe Interact.* **11**, 523-529.

Eshdat, Y., Holland, D., Faltin, Z., and Ben-Hayyim, G. (1997). Plant gluathione peroxidases. *Physiol. Plant.* **100**, 234 - 240.

Ewing, B., Hillier, L., Wendl, M. C., and Green, P. (1998). Base-calling of automated sequencer traces using *Phred*: I. Accuracy assessment. *Genome Res.* **8**, 175 - 185.

Fambrough, *et al.* (1999). Diverse signalling pathways activated by growth factor receptors induce broadly overlapping, rather than indiependent, sets of genes. *Cell* 97: 727-741.

Feinberg, A. P., and Vogelstein, B. (1983). A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Analytical Biochemistry* **132**, 6-13.

Feldmann, K. A., Malmberg, R. L., and Dean, C. (1994). Mutagenesis in *Arabidopsis*. In *Arabidopsis*, C. R. Somerville and E. M. Meyerowitz, eds. (New York: Cold Spring Harbor Laboratory Press), pp. 137 - 172.

Felsenstein, J. (1996). Inferring phylogenies from protein sequences by parsimony, distance, and likelihood methods. *Methods Enzymol.* **266**, 418 - 427.

Flor, H. (1956). The complementary genetic systems in flax and flax rust. *Adv. Genet.* **8**, 29-54.

Fodor, J., Gullner, G., Adam, A. L., Barna, B., Komives, T., and Kiraly, Z. (1997). Local and systemic responses of antioxidants to tobacco mosaic virus infection and to salicylic acid in tobacco. *Plant Physiol.* **114**, 1443 - 1451.

Friedrich, L., Lawton, K., Ruess, W., Masner, P., Specker, N., Gut-Ruella, M., Meier, B., Dincher, S., Staub, T., Uknes, S., Metraux, J.-P., Kessmann, H., and Ryals, J. (1996). A benzothiadiazole derivate induces systemic acquired resistance in tobacco. *Plant J.* **10**, 61 - 70.

Frugoli, J. A., Zhong, H. H., Nuccio, M. L., McCourt, P., McPeek, M. A., Thomas, T. L., and McClung, C. R. (1996). Catalase is encoded by a multigene family in *Arabidopsis thaliana* (L.) Heynh. *Plant Physiol.* **112**, 327 - 336.

Frye, C. A., and Innes, R. W. (1998). An *Arabidopsis* mutant with enhanced resistance to powdery mildew. *Plant Cell* **10**, 947 - 956.

Gaffney, T., Friedrich, L., Vernooij, B., Negrotto, D., Nye, G., Uknes, S., Ward, E., Kessmann, H., and Ryals, J. (1993). Requirement of salicylic acid for the induction of systemic acquired resistance. *Science* **261**, 754-756.

Gehring, W. J. (1987). Homeoboxes in the study of development. *Science* **236**, 1245 - 1252.

Gerhold, *et al.* (1999). DNA chips: promising toys have become powerful tools. TIBS 24-May 168-173.

Georgopoulos, S. G. (1995). The genetics of fungicide resistance. In Modern selective fungicides, H. Lyr, ed. (Jena: Fischer), pp. 39 - 52.

Giese, H., Hippe-Sanwald, S., Somerville, S., and Weller, J. (1997). *Erysiphe graminis*. In The Mycota, Carroll and Tudzynski, eds. (Berlin: Springer Verlag), pp. 55 - 78.

Gish, W., and States, D. J. (1993). Identification of protein coding regions by database similarity search. *Nat. Genetics* **3**, 266 - 272.

Glazebrook, J., and Ausubel, F. M. (1994). Isolation of phytoalexin-deficient mutants of *Arabidopsis thaliana* and characterization of their interactions with bacterial pathogens. *Proc. Natl. Acad. Sci. USA* **91**, 8955 - 8959.

Glazebrook, J., Rogers, E. E., and Ausubel, F. M. (1996). Isolation of *Arabidopsis* mutants with enhanced disease susceptibility by direct screening. *Genetics* **143**, 973 - 982.

Glazebrook, J., Rogers, E. E., and Ausubel, F. M. (1997). Use of *Arabidopsis* for genetic dissection of plant defense responses. *Annu. Rev. Genetics* **31**, 547 - 569.

Glazebrook, J., Zook, M., Mert, F., Kagan, I., Rogeers, E. E., Crute, I. R., Holub, E. B., Hammerschmidt, R., and Ausubel, F. M. (1997). Phytoalexin-deficient mutants of *Arabidopsis* reveal that *PAD4* encodes a regulatory factor and that four *PAD* genes contribute to downy mildew resistance. *Genetics* **146**, 381 - 392.

Goodman, H. M., Ecker, J. R., and Dean, C. (1995). The genome of *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. USA* **92**, 10831 - 10835.

Gordon, D., Abajian, C., and Green, P. (1998). *Consed:* A graphical tool for sequence finishing. *Genome Res.* **8**, 195 - 202.

Görlach, J., Raesecke, H.-R., Rentsch, D., Regenass, M., Roy, P., Zala, M., Keel, C., Boller, T., Amrhein, N., and Schmid, J. (1995). Temporally distinct accumulation of transcripts encoding enzymes of the prechorismate pathway in elicitor-treated, cultured tomato cells. *Proc. Natl. Acad. Sci. USA* **92**, 3166 - 3170.

Görlach, J., Volrath, S., Knauf-Breiter, G., Hengy, G., Beckhove, U., Kogel, K.-H., Oostendorp, M., Staub, T., Ward, E., Kessmann, H., and Ryals, J. (1996). Benzothiadiazole, a novel class of inducers of systemic acquired resistance, activates gene expression and disease resistance in wheat. *Plant Cell* **8**, 629 - 643.

Gosti, F., Berauche, N., Vartanian, N., and Giraudat, J. (1995). Abscisic acid-dependent and -independent regulation of gene expression by progressive drought in *Arabidopsis thaliana*. *Mol. Gen Genet.* **246**, 10 - 18.

Grant, M. R., Godiard, L., Straube, E., Ashfield, T., Lewald, J., Sattler, A., Innes, R. W., and Dangl, J. L. (1995). Structure of the *Arabidopsis RPM1* gene enabling dual specificity disease resistance. *Science* **269**, 843-846.

Gray, J., Close, P. S., Briggs, S. P., and Johal, G. S. (1997). A novel suppressor of cell death in plants encoded by the *Lls1* gene of Maize. *Cell* **89**, 25 - 31.

Greenberg, J. T., and Ausubel, F. M. (1993). *Arabidopsis* mutants compromised for the control of cellular damage during pathogenesis and aging. *Plant J.* **4**, 327-342.

Gus-Mayer, S., Naton, B., Hahlbrock, K., and Schmelzer, E. (1998). Local mechanical stimulation induces components of the pathogen defense response in parsley. *Proc. Natl. Acad. Sci. USA* **95**, 8398-8403.

Guyer, D., Patton, D., and Ward, E. (1995). Evidence for cross-pathway regulation of metabolic gene expression in plants. *Proc. Natl. Acad. Sci. USA* **92**, 4997 - 5000.

Hahlbrock, K., and Scheel, D. (1989). Physiology and molecular biology of phenypropanoid metabolism. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **40**, 347-369.

Hahlbrock, K., Scheel, D., Logemann, E., Nuernberger, T., Parniske, M., Reinhold, S., Sacks, W. R., and Schmelzer, E. (1995). Oligopeptide elicitor-mediated defense gene activation in cultured parsley cells. *Proc. Natl. Acad. Sci. USA* **92**, 4150 - 4157.

Halliwell, B., and Gutteridge, J. M. C. (1989). *Free radicals in biology and medicine* (Oxford: Clarendon Press).

Hanahan, D. (1983). Sudies of transformation of *Escherichia coli* with plasmids. *J. Mol. Biol.* **166**, 557 - 580.

Hancock, R. E. W., Falla, T., and Brown, M. (1995). Cationic bactericidal peptides. *Adv. Microbial Physiol.* **37**, 135 - 175.

Harlow, E., and Lane, D. (1988). *Antibodies - a laboratory manual* (Cold Spring Harbor: Cold Spring Harbor Laboratory Press).

Haughn, G., and Somerville, C. R. (1987). Selection for herbicide resistance at the whole-plant level. In Biotechnology in agricultural chemistry: American Chemical Society), pp. 98-107.

Hauser, M. T., Adhami, F., Dorner, M., Fuchs, E., and Gloessl, J. (1998). Generation of co-dominant PCR-based markers by homo- and heteroduplex analysis on high resolution gels. *Plant J.* **16**, 117 - 125.

Hoffman, C. S., and Winston, F. (1987). A ten-minute DNA preparation from yeast efficiently releases autonomous plasmids for transformation of *Escherichia coli*. *Gene* **57**, 267 - 272.

Hofte, H., Desprez, T., Amselem, J., Chiapello, H., Caboche, M., Moisan, A., Jourjon, M.-F., Caharpenteau, L., Berthomieu, P., Guerrier, D., Giraudat, J., Quigley, F., Thomas, F., Yu, D.-Y., Mache, R., Raynal, M., Cooke, R., Grellet, F., Delseny, M., Parmentier, Y., de Marcillac, G., Gigot, C., Fleck, J., Phillips, J., Axelos, M., Bardet, C., Tremousaygue, D., and Lescure, B. (1993). An inventory of 1152 expressed sequence tags obtained by partial sequencing of cDNAs from *Arabidopsis thaliana*. *Plant J.* **4**, 1051-1061.

Holub, E. B., and Beynon, J. L. (1997). Symbiology of Mouse-Ear Cress (*Arabidopsis thaliana*) and Oomycetes. *Advances in Botanical Research, Advances in Plant Pathology* **24**, 227 - 273.

Holub, E. B., Beynon, J. L., and Crute, I. R. (1994). Phenotypic and genotypic characterization of interactions between isolates of *Peronospora parasitica* and accessions of *Arabidopsis thaliana*. *Mol. Plant Microbe Interact.* 7, 223 - 239.

Hu, G., Richter, T. E., Hulbert, S. H., and Pryor, T. (1996). Disease lesion mimicry caused by mutations in the rust resistance gene *rp1*. *Plant Cell* **8**, 1367 - 1376.

Hunt, M. D., Delaney, T. P., Dietrich, R. A., Weymann, K. B., Dangl, J. L., and Ryals, J. A. (1997). Salicylate-independent lesion formation in *Arabidopsis lsd* mutants. *Mol. Plant Microbe Interact.* **10**, 531 - 536.

Iyer, *et al.* (1999). The transcriptional program in the response or human fibroblasts to serum. *Science* 283: 83-87.

Jabs, T., Dietrich, R., and Dangl, J. (1996). Extracellular superoxide is necessary and sufficient for runaway cell death in an *Arabidopsis* mutant. *Science* **273**, 1853 - 1856.

Jones, J. D. G. (1996). Plant disease resistance genes: structure, function and evolution. *Curr. Opin. Biotechnol.* **7**, 155 - 160.

Kamalay, J. C., and Goldberg, R. B. (1980). Regulation of structural gene expression in tobacco. *Cell* **19**, 935 - 946.

Kauschmann, A., Jessop, A., Koncz, C., Szekeres, M., Willmitzer, L., and Altmann, T. (1996). Genetic evidence for an essential role of brassinosteroids in plant development. *Plant J.* **9**, 701 - 713.

Keller, H., Blein, J-P., Bonnet, P., and Ricci, P. (1996). Physiological and molecular characteristics of elicitin-induced systemic acquired resistance in tobacco. *Plant Physiol.* **110**, 365 - 376.

Kehoe, *et al.* (1999). DNA microarrays for studies of higher plants and other photosynthetic organisms. *Trends in Plt. Sci.* 4: 38-41.

Keogh, R. C., Deverall, B. J., and McLeod, S. (1980). Comparison of histological and physiological responses to *Phakopsora pachyrhizi* in resistant and susceptible soybean. *Trans. Br. Mycol. Soc.* **74**, 329-333.

Kieber, J. J. (1997). The ethylene response pathway in *Arabidopsis*. *Annu. Rev. Plant. Physiol. Plant Mol. Biol.* **48**, 277 - 296.

Kim, U. J., Birren, B. W., Slepak, T., Mancino, V., Boysen, C., Kang, H. L., Simon, M. I., and Shizuya, H. (1996). Construction and characterization of a human bacterial artificial chromosme library. *Genomics* **34**, 213 - 218.

Kimmel, B., Palozzolo, M. J., Martin, C. H., Boeke, J. D., and Devine, S. E. (1997). Transposon-mediated DNA sequencing. In Genome Analysis, a laboratory manual, B. Birren, E. Green, P. Hieter and R. Myers, eds. (New York: Cold Spring Harbor Laboratory Press), pp. 455 - 532.

Klee, H. J., Muskopf, Y. M., and Gasser, C. S. (1987). Cloning of an *Arabidopsis thaliana* gene encoding 5-*enol*pyruvylshikimate-3-phosphate synthase: sequence analysis and manipulation to obtain glyphosate-tolerant plants. *Mol. Gen. Genet.* **210**, 437 - 442.

Knoester, M., van Loon, L. C., van den Heuvel, J., Hennig, J., Bol, J. F., and Linthorst, H. J. M. (1998). Ethylene-insensitive tobacco lacks nonhost resistance against soil-borne fungi. *Proc. Natl. Acad. Sci. USA* **95**, 1933-1937.

Koch, E., and Slusarenko, A. (1990). *Arabidopsis* is susceptible to infection by a downy mildew fungus. *Plant Cell* **2**, 437-445.

Kombrink, E., and Hahlbrock, K. (1986). Resonses of cultured parsley cells to elicitors from phytopathogenic fungi. *Plant Physiol.* **81**, 216 - 221.

Kombrink, E., and Somssich, I. (1997). Pathogenesis-related proteins and plant defense. In The Mycota, Carroll and Tudzynski, eds. (Berlin: Springer Verlag), pp. 107 - 125.

Koncz, C., and Schell, J. (1986). The promoter of TL-DNA gene 5 controls the tissue-specific expression of chimaeric genes carried by a novel type of *Agrobacterium* binary vector. *Mol. Gen. Genet.* **204**, 383 - 396.

Konieczny, A., and Ausubel, F. M. (1993). A procedure for mapping *Arabidopsis* mutations using co-dominant ecotype-specific PCR-based markers. *Plant J.* **4**, 403-410.

Koorneef, M., and Stam, P. (1992). Genetic Analysis. In Methods in *Arabidopsis* Research, C. Koncz, N.-H. Chua and J. Schell, eds. (Singapore: World Scientific), pp. 83 - 99.

Kosambi, D. D. (1944). The estimation of map distances from recombination values. *Ann. Eugen.* **12**, 172-175.

Kubo, A., Saji H., Tanaka, K., and Kondo, N. (1992). Cloning and sequencing of a cDNA encoding ascorbate peroxidase from *Arabidopsis thaliana*. *Plant Mol. Biol.* **18**, 691 - 701.

Lagrimini, L. M., Burkhart, W., Moyer, M., and Rothstein, S. (1987). Molecular cloning of complementary DNA encoding the lignin-forming peroxidase from tobacco: molecular analysis and tissue-specific expression. *Proc. Natl. Acad. Sci. USA* **84**, 7542-7546.

Landegren, U. (1999). A prelude to the stars: genomics in Cannes. *Trends Biotech.* **17**, 1 - 3.

Lander, E. S., Green, P., Abrahamson, J., Barlow, A., Daly, M. J., Lincoln, S. E., and Newburg, L. (1987). MAPMAKER: An interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. *Genomics* **1**, 174-181.

Laugé, R., Joosten, M. A. A. J., Haanstra, J. P. W., Goodwin, P. H., Lindhout, P., and De Wit, P. J. G. M. (1998). Successful search for a resistance gene in tomato targeted against a virulence factor of a fungal pathogen. *Proc. Natl. Acad. Sci. USA* **95**, 9014-9018.

Lawton, K., Friedrich, L., Hunt, M., Weymann, K., Delaney, T., Kessmann, H., Staub, T., and Ryals, J. (1996). Benzothiadiazole induces disease resistance in *Arabidopsis* by activation of the systemic acquired resistance signal transduction pathway. *Plant J.* **10**, 71 - 82.

Lawton, K., Potter, S., Uknes, S., and Ryals, J. (1994). Systemic acquired resistance signal transduction in *Arabidopsis* is ethylene independent. *Plant Cell* **6**, 581-588.

Lebel, E., Heifetz, P., Thorne, L., Uknes, S., Ryals, J., and Ward, E. (1998). Functional analysis of regulatory sequences controlling PR-1 gene expression in *Arabidopsis*. *Plant J.* **16**, 223 - 233.

Lemieux, *et al.* (1998). Overview of DNA chip technology. *Molecular Breeding* 4: 277-289.

Lermontova, I., Kruse, E., Mock, H. P., and Grimm, B. (1997). Cloning and characterization of a plastidal and a mitochondrial isoform of tobacco protoporphyrinogen IX oxidase. *Proc. Natl. Acad. Sci USA* **94**, 8895 - 8900.

Li, S. L., and Redei, G. P. (1969). Thiamine mutants of the crucifer *Arabidopsis*. *Biochem. Genet.* **3**, 163 - 170.

Lincoln, S., Daly, M., and Lander, E. (1992). Constructing genetic maps with Matchmaker/EXP 3.0: Whitehead Institute Technical Report).

Liu, Y. G., Mitsukawa, N., Lister, C., Dean, C., and Whittier, R. F. (1996). Isolation and mapping of a new set of RFLP markers in *Arabidopsis thaliana* using recombinant inbred lines. *Plant J.* **10**, 733 - 736.

Liu, Y.-G., Mitsukawa, N., Oosumi, T., and Whittier, R. (1995). Efficient isolation and mapping of *Arabidopsis thaliana* T-DNA insert junctions by thermal asymmetric interlaced PCR. *Plant J.* **8**, 457 - 463.

Low, P. S., and Merida, J. R. (1996). The oxidative burst in plant defense: Function and signal transduction. *Physiol. Plant.* **96**, 533 - 542.

Lund, S. T., Stall, R. E., and Klee, H. J. (1998). Ethylene regulates the susceptible response to pathogen infection in tomato. *Plant Cell* **10**, 371 - 382.

Maleck, K. and Dietrich, R. A. (1999). Synergies and antagonisms between plant pathogen defense systems. *Trends Pl. Sci., in press*

Maleck, K., and Lawton, K. (1998). Plant strategies for resistance to pathogens. *Curr. Opin. Biotechnol.* **9**, 208-213.

Marrs, K. A. (1996). The functions and reguation of glutathione S-transferases in plants. *Ann. Rev. Plant Physiol. Plant Mol. Biol.* **47**, 127 - 158.

Martin, G. B., Frary, A., Wu, T., Brommonschenkel, S., Chunwongse, J., Earle, E. D., and Tanksley, S. D. (1994). A member of the tomato *Pto* gene family confers sensitivity to Fenthion resulting in rapid cell death. *Plant Cell* **6**, 1543-1552.

Mauch-Mani, B., and Slusarenko, A. J. (1996). Production of salicylic acid precursors is a major function of phenylalanine ammonia-lyase in the resistance of *Arabidopsis* to *Peronospora parasitica*. *Plant Cell* **8**, 203 - 212.

McAdams, H., and Shapiro, L. (1995). Circuit simulation of genetic networks. *Science* **269**, 650 - 656.

McCall, K., and Steller, H. (1997). Facing death in the fly: genetic analysis of apoptosis in *Drosophila*. *Trends Genet.* **13**, 222 - 226.

McConn, M., Creelman, R. A., Bell, E., Mullet, J. E., and Browse, J. (1997). Jasmonate is essential for insect defense in *Arabidopsis*. *Proc. Natl. Acad. Sci. USA* **94**, 5473 - 5477.

McDowell, J. M., Dhandaydham, M., Long, T. A., Aarts, M. G. M., Goff, S., Holub, E. B., and Dangl, J. L. (1998). Intragenic recombination and diversifying selection contribute to the evolution of downy mildew resistance at the *RPP8* locus of *Arabidopsis*. *Plant Cell* **10**, 1861 - 1874.

Mednik, I. G. (1988). On methods evaluating the frequencies of induced mutations in *Arabidopsis* based on embryo-test data. *Arabidopsis Information Service* **26**, 67 - 72.

Melan, M., A.,, Dong, X., Endara, M. E., Davis, K. R., Ausubel, F. M., and Peterman, T. K. (1993). An *Arabidopsis thaliana* lipoxygenase gene can be induced by pathogens, abscisic acid, and methyl jasmonate. *Plant Physiol.* **101**, 441 - 450.

Melnick, M. B., Noll, E., and Perrimon, N. (1993). The Drosophila *stubarista* phenotype is associated with a dosage effect of the putative ribosome-associated protein D-p40 on *spineless*. *Genetics* **135**, 553 - 564.

Memelink, J., Swords, K. M. M., de Kam, R. J., Schilperoort, R. A., Hoge, J. H. C., and Staehelin, L. A. (1993). Structure and regulation of tobacco extensins. *Plant J.* 4, 1011 - 1022.

Merlot, S., and Giraudat, J. (1997). Genetic analysis of abscisic acid signal transduction. *Plant Physiol.* **114**, 751 - 757.

Metraux, J.-P., Burkhart, W., Moyer, M., Dincher, S., Middlesteadt, W., Williams, S., Payne, G., Carnes, M., and Ryals, J. (1989). Isolation of complementary DNA encoding a chitinase with structural homology to a bifunctional lysozyme/chitinase. *Proc. Natl. Acad. Sci. USA.* **86**, 896-900.

Meyerowitz, E. M. (1989). *Arabidopsis*, a useful weed. *Cell* **56**, 263-269.

Meyerowitz, E. M., and Somerville, C. R. (1994). *Arabidopsis* (New York: Cold Spring Harbor Laboratory Press).

Michelmore, R. W., and Meyers, B. C. (1998). Clusters of Resistance genes in plants evolve by divergent selection and a birth-and-death process. *Genome Res.* **8**, 1113 - 1130.

Millar, A. J., Short, S. R., Hiratsuka, K., Chua, N.-H., and Kay, S. (1992). Firefly luciferase as a reporter of regulated gene expression in higher plants. *Plant Mol. Biol. Rep.* **10**, 324 - 337.

Mindrinos, M., Katagiri, F., Yu, G.-L., and Ausubel, F. M. (1994). The *A. thaliana* disease resistance gene *RPS2* encodes a protein containing a nucleotide-binding site and leucine-rich repeats. *Cell* **78**, 1089-1099.

Molina, A., and Garcia-Olmedo, F. (1997). Enhanced tolerance to bacterial pathogens caused by the transgenic expression of barley lipid transfer protein LTP2. *Plant J.* **12**, 669 - 675.

Mozo, T., Fischer, S., Meier-Ewert, S., Lehrach, H., and Altmann, T. (1998). Use of the IGF BAC library for physical mapping of the *Arabidopsis thaliana* genome. *Plant J.* **16**, 377 - 384.

Mozo, T., Fischer, S., Shizuya, H., and Altmann, T. (1998). Construction and characterization of the IGF *Arabidopsis* BAC library. *Mol. Gen. Genet.* **258**, 562 - 570.

Mundy, J., Mayer, R., and Chua, N.-H. (1995). Cloning genomic sequences using long-range PCR. *Plant Mol. Biol. Rep.* **13**, 156 - 163.

Murashige, T., and Skoog, F. (1962). A revised medium for rapid growth and bioassays with tobacco tissue cultures. *Physiol. Plant.* **15**, 473-497.

Myers, R. M., Larin, Z., and Maniatis, T. (1985). Detection of single base pair substitutions by ribonuclease cleavage at msmatches in RNA:DNA duplexes. *Science* **230**, 1242 - 1246.

Neff, M. M., Neff, J. D., Chory, J., and Pepper, A. E. (1998). dCAPS, a simple technique for the genetic analysis of single nucleotide polymorphisms: experimental applications in *Arabidopsis thaliana* genetics. *Plant J.* **14**, 387-392.

Newman, T., de Bruijn, F. J., Green, P., Keegstra, K., Kende, H., McIntosh, L., Ohlrogge, J., Raikhel, N., Somerville, S., Thomashow, M., Retzel, E., and Somerville, C. (1994). Genes Galore: A summary of methods for accessing results of large-scale partial sequencing of anonymous *Arabidopsis* cDNA clones. *Plant Physiol.* **106**, 1241 - 1255.

Nizetic, D., Drmanac, R., and Lehrach, H. (1991). An improved bacterial colony lysis procedure enables direct DNA hybridization using short (10,11 bases) oligonucleotides to cosmids. *Nucl. Acid. Res.* **19**, 182.

Oerke, E.-C., Dehne, H.-W., Schönbeck, F., and Weber, A. (1994). *Crop Production and Crop Protection* (Amsterdam: Elsevier Science B. V.).

O'Farrell, P. H. (1975). High resolution two-dimensional electrophoresis of proteins. *J. Biol. Chem.* **250**, 4007 - 4021.

Ogas, J., Cheng, J.-C., Sung, Z. R., and Somerville, C. (1997). Cellular differentiation regulated by gibberellin in the *Arabidopsis thaliana pickle* mutant. *Science* **277**, 91 - 94.

Ow, D., Wood, K. V., DeLuca, M., DeWet, J. R., Helinski, D. R., and Howell, S. H. (1986). Transient and stable expression of the firefly luciferase gene in plant cells and transgenic plants. *Science* **234**, 856 - 859.

Parker, J. E., Holub, E. B., Frost, L. N., Falk, A., Gunn, N. D., and Daniels, M. J. (1996). Characterization of *eds1*, a mutation in *Arabidopsis* suppressing resistance to *Peronospora parasitica* specified by several different *RPP* Genes. *Plant Cell* **8**, 2033 - 2046.

Patton, D. A., Johnson, M., and Ward, E. R. (1996). Biotin synthase from *Arabidopsis thaliana*. *Plant Physiol.* **112**, 371 - 378.

Paxton, J. D. (1981). Phytoalexins- A working redefinition. *Pytopathol. Z.* **101**, 106 - 109.

Penninckx, I. A. M. A., Eggermont, K., Terras, F. R. G., Thomma, B. P. H. J., De Samblanx, G. W., Buchala, A., Metraux, J.-P., Manners, J. M., and Broekaert, W. F. (1996). Pathogen-induced systemic activation of a plant defensin gene in *Arabidopsis* follows a salicylic acid-independent pathway. *Plant Cell* **8**, 2309 - 2323.

Penninckx, I. A. M. A., Thomma, B. P. H. J., Buchala, A., Metraux, J. P., and Broekaert, W. F. (1998). Concomiant activation of jasmonate and ethylene response pathways is required for induction of a plant defensin gene in *Arabidopsis*. *Plant Cell* **10**, 2103 - 2113.

Pennington, S. R., Wilkins, M. R., Hochstrasser, D. F., and Dunn, M. J. (1997). Proteome analysis: from protein characterization to biological function. *Trends Cell Biol.* 7, 168 - 173.

Peterhänsel, C., Freialdenhofen, A., Kurth, J., Kolsch, R., and Schulze-Lefert, P. (1997). Interaction analysis of genes required for resistance responses to powdery mildew in barley reveal distinct pathways leading to leaf cell death. *Plant Cell* **9**, 1397 - 1409.

Peterman, T. K., and Goodmann, H. M. (1991). The glutamine synthetase gene family of *Arabidopsis thaliana*: light-regulation and differential expression in leaves, roots and seeds. *Mol. Gen. Genet.* **230**, 145 - 154.

Pieterse, C. M. J., van Wees, S. C. M., Hoffland, E., van Pelt, J. A., and van Loon, L. C. (1996). Systemic resistance in *Arabidopsis* induced by biocontrol bacteria is independent of salicylic acid accumulation and pathogenesis-related gene expression. *Plant Cell* **8**, 1225 - 1237.

Pieterse, C. M. J., van Wees, S. C. M., van Pelt, J. A., Knoester, M., Laan, R., Gerrits, H., Weisbeek, P. J., and van Loon, L. C. (1998). A novel signaling pathway controlling induced systemic resistance in *Arabidopsis*. *Plant Cell* **10**, 1571-1580.

Plesch, G., Stoermann, K., Torres, J. T., Walden, R., and Somssich, I. E. (1997). Developmental and auxin-induced expression of the *Arabidopsis prha* homeobox gene. *Plant J.* **12**, 635 - 647.

Potter, S., Uknes, S., Lawton, K., Ward, E., and Ryals, J. (1993). Regulation of a hevein-like gene in *Arabidopsis*. *Mol. Plant Microbe Interact.* **6**, 680-685.

Price, R. A., Palmer, J. D., and Al-Shehbaz, I. A. (1994). Systematic relationships of *Arabidopsis*: a molecular and morphological perspective. In *Arabidopsis*, E. M. Meyerowitz and C. R. Somerville, eds.: Cold Spring Harbor Laboratory Press).

Pruitt, R. E., and Meyerowitz, E. M. (1986). Characterization of the genome of *Arabidopsis thaliana*. *J. Mol. Biol.* **187**, 169 - 183.

Ptashne, M. (1992). *A genetic switch: Phage l and higher organisms*, 2nd Edition (Cambridge, MA: Blackwell Scientific Publications).

Rafalski, J. A., Vogel, J. M., Morgante, M., Powell, W., Andre, C., and Tingey, S. V. (1996). Generating and using DNA markers in plants. In Nonmammalian Genomic Analysis, B. Birren and E. Lai, eds. (San Diego: Academic Press), pp. 75 - 134.

Rance, I., Fournier, J., and Esquerre-Taugaye, M.-T. (1998). The incompatible interaction between *Phytophthora parasitica* var. *nicotianae* race 0 and tobacco is suppressed in transgenic plants expressing antisense lipoxygenase sequences. *Proc. Natl. Acad. Sci. USA* **95**, 6554-6559.

Rasmussen, J. B., Hammerschmidt, R., and Zook, M. N. (1991). Systemic induction of salicylic acid accumulation in cucumber after inoculation with *Pseudomonas syringae* pv *syringae*. *Plant Physiol.* **97**, 1342-1347.

Redei, G. P. (1975). *Arabidopsis* as a genetic tool. *Annu. Rev. Genetics* **9**, 111 - 127.

Redei, G. P., and Koncz, C. (1992). Classical mutagenesis. In Methods in *Arabidopsis* research, C. Koncz, N.-H. Chua and J. Schell, eds. (Singapore: World Scientific).

Reinold, S., and Hahlbrock, K. (1996). Biphasic Temporal and Spatial Induction Patterns of Defense-Related mRNAs and Proteins in Fungus-Infected Parsley Leaves. *Plant Physiol.* **112**, 131 - 140.

Reiter, R. S., Williams, J. G. K., Feldman, K. A., Rafalski, J. A., Tingey, S. V., and Scolnik, P. A. (1992). Global and local genome mapping in *Arabidopsis thaliana* by using recombinant inbred lines and random amplified polymorphic DNAs. *Proc. Natl. Acad. Sci.USA* **89**, 1477-1481.

Reuber, T. L., and Ausubel, F. M. (1996). Isolation of *Arabidopsis* genes that differentiate between resistance responses mediated by the *RPS2* and *RPM1* disease resistance genes. *Plant Cell* **8**, 241 - 249.

Reuber, T. L., Plotnikova, J. M., Dewdney, J., Rogers, E. E., Wood, W., and Ausubel, F. M. (1998). Correlation of defense gene induction defects with powdery mildew susceptibility in *Arabidopsis* enhanced disease susceptibility mutants. *Plant J.* **16**, 473 - 485.

Reymond, P., and Farmer, E. E. (1998). Jasmonate and salicylate as global signals for defense gene expression. *Curr. Opin. Plant Biology* **1**, 404 - 411.

Ritter, C., and Dangl, J. L. (1996). Interference between two specific pathogen recognition events mediated by distinct plant disease resistance genes. *Plant Cell* **8**, 251 - 257.

Rogers, S. O., and Bendich, A. J. (1988). Extraction of DNA from plant tissues. In Plant Molecular Biology Manual, S. B. Gelvin and R. A. Schilperpoort, eds. (Dordrecht: Kluwer Academic Publishers), pp. A6-1 - A6-11.

Rose, A. B., Casselman, A. L., and Last, R. L. (1992). A phosphoribosylanthranilate transferase gene is defective in blue flourescent *Arabidopsis thaliana* tryptophan mutants. *Plant Physiol.* **100**, 582 - 592.

Ross, A. F. (1961a). Localized acquired resistance to plant virus infection in hypersensitive hosts. *Virology* **14**, 329-339.

Ross, A. F. (1961b). Systemic acquired resistance induced by localized virus infections in plants. *Virology* **14**, 340-358.

Roth, *et al.* (1998). Finding DNA regulatory motifs within unaligned noncoding sequences clustered by whole-genome mRNA quantitation. *Nature Biotech* 16, 939-945.

Ruan, Y., Gilmore, J., and Conner, T. (1998). Towards *Arabidopsis* genome analysis: monitoring expression profiles of 1400 genes using cDNA microarrays. *Plant J.* **15**, 821 - 833.

Rushton, P. J., and Somssich, I. E. (1998). Transcriptional control of plant genes responsive to pathogens. *Curr. Opin. Pl. Biol.* **1**, 311 - 315.

Ryals, J., Weymann, K., Lawton, K., Friedrich, L., Ellis, D., Steiner, H.-Y., Johnson, J., Delaney, T. P., Jesse, T., Vos, P., and Uknes, S. (1997). The *Arabidopsis NIM1* protein shows homology to the mammalian transcription factor inhibitor IkB. *Plant Cell* **9**, 425 - 439.

Ryan, C. (1990). Protease inhibitors in plants: Genes for improving defenses against insects and pathogens. *Ann. Rev. Phytopathol.* **28**, 425-49.

Salmeron, J. M., Oldroyd, G. E. D., Rommens, C. M. T., Scofield, S. R., Kim, H.-S., Lavelle, D. T., Dahlbeck, D., and Staskawicz, B. J. (1996). Tomato *Prf* is a member of the leucine-rich repeat class of plant disease resistance genes and lies embedded within the *Pto* kinase gene cluster. *Cell* **86**, 123-133.

Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). *Molecular Cloning*, 2nd edition Edition (New York: Cold Spring Harbor Laboratory Press).

Sanger, F., Nicklen, S., and Coulson, A. R. (1977). DNA sequncing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**, 5463 - 5467.

Santoni, V., Bellini, C., and Caboche, M. (1994). Use of two-dimensional protein-pattern analysis for the characterization of *Arabidopsis thaliana* mutants. *Planta* **192**, 557 - 566.

Schena, *et al.* (1998). Microarrays: biotecnology's discovery platform for functional genomics. *TIBTECH* 16:301-306.

Schena, M., Shalon, D., Davis, R., and Brown, P. (1996). Parallel human genome analysis: microarray-based expression monitoring of 1000 genes. *Proc. Natl. Acad. Sci. USA* **93**, 10614 - 10619.

Schena, M., Shalon, D., Davis, R. W., and Brown, P. O. (1995). Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* **270**, 467 - 470.

Schmidt, R. (1998). The *Arabidopsis thaliana* genome: towards a complete physical map. In *Arabidopsis*, M. Anderson and J. Roberts, eds. (Boca Raton: Sheffield Academic Press), pp. 1 - 30.

Schott, E. J., Davis, K. R., Dong, X., Mindrinos, M., Guevara, P., and Ausubel, F. M. (1990). *Pseudomonas syringae* infection of *Arabidopsis thaliana* as a model system for studying plant-bacterial interactions. In *Pseudomonas*: biotransformations, pathogenesis, and evolving biotechnology, S. Siver, A. Chakrabarty and B. Iglewski, eds. (Washington, D.C.: American Society for Microbiology), pp. 82-90.

Schulz, B., Bennett, M. J., Dilkes, B. P., and Feldmann, K. A. (1994). T-DNA tagging in *Arabidopsis thaliana*: cloning by gene disruption. In Plant Molecular Biology Manual (Belgium: Kluwer Academic Press).

Shah, J., Tsui, F., and Klessig, D. F. (1997). Characterization of a salicylic acid-insenstivie mutant (*sail*) of *Arabidopsis thaliana*, identified in a selective screen utilizing the SA-inducible expression of the *tms2* gene. *Mol. Plant Microbe Interact.* **10**, 69 - 78..

Shalon, D., Smith, S. J., and Brown, P. O. (1996). A DNA microarray system for analyzing complex DNA samples using two-color fluorescent probe hybridization. *Genome Res.* **6**, 639 - 645.

Shirley, B. W., Kubasek, W. L., Storz, G., Bruggemann, E., Koorneef, M., Ausubel, F. M., and Goodman, H. M. (1995). Analysis of *Arabidopsis* mutants deficient in flavonoid biosynthesis. *Plant J.* **8**, 659 - 671.

Short, J. M., Fernandez, J. M., Sorge, J. A., and Huse, W. D. (1988). l Zap: a bacteriophage l expression vector with *in vivo* excision properties. *Nucl. Acid Res.* **16**, 7583 - 7587.

Simms, E. L. (1996). The evolutionary genetics of plant-pathogen systems. *BioScience* **46**, 136 - 145.

Sinnett, D., Richer, C., and Baccichet, A. (1998). Isolation of stable Bacterial Artificial Chromosome DNA using a modified alkaline lysis methods. *Biotechniques* **24**, 752 - 754.

Somssich, I., and Hahlbrock, K. (1998). Pathogen defence in plants-a paradigm of biological complexity. *Trends Pl. Sci.* **3**, 86 - 90.

Sonnewald, U., Brauer, M., von Schaewen, A., Stitt, M., and Willmitzer, L. (1991). Transgenic tobacco plants expressing yeast-derived invertase in either the cytosol, vacuole or apoplast: a powerful tool for studying sucrose metabolism and sink/source interactions. *Plant J.* **1**, 95 - 106.

Sonstad, D. D., De Luca, V., Brisson, N., Kurz, W. G. W., and Nessler, C. L. (1990). High levles of tryptamine accumulation in transgenic tobacco expressing tryptophan decarboxylase. *Plant Physiol.* **94**, 1410 - 1413.

Staswick, P. E., Yuen, G. Y., and Lehman, C. C. (1998). Jamonate signaling mutants of *Arabidopsis* are susceptible to the soil fungus *Pythium irregulare*. *Plant J.* **15**, 747-754.

Stiekema, W. J., and Pereira, A. (1998). Unravelling the genome by genome sequencing and function analysis. In *Arabidopsis*, M. Anderson and J. Roberts, eds. (Boca Raton: Sheffield Academic Press), pp. 31 - 62.

Stitt, M., von Schaewen, A., and Willmitzer, L. (1990). "Sink" regulation of photosynthetic metabolism in transgenic tobacco plants expressing yeast invertase in their cell wall involves a decrease of the Calvin-cycle enzymes and an increase of glycolytic enzymes. *Planta* **183**, 40 - 50.

Sugimoto, A., Hozak, R. R., Nakashima, T., Nishimoto, T., and Rothman, J. H. (1995). *dad-1*, an endogenous programmed cell death suppressor in *Caenorhabditis elegans* and vertebrates. *EMBO J.* **14**, 4434 - 4441.

Switzer, R. C., Merril, C. R., and Shifrin, S. (1979). A highly sensitive stain for detecting proteins and peptides in polyacrylamide gels. *Anal. Biochem.* **98**, 231 - 237.

Tavozoie, *et al.* (1999). Systematic determination of genetic network architecture. *Nature Genetics* 22, 281-285.

Terras, F., R., G.,, Eggermont, K., Kovaleva, V., Raikhel, N. V., Osborn, R. W., Kester, A., Rees, S. B., Torrekens, S., Van Leuven, F., Vanderleyden, J., Cammue, B. P. A., and Broekaert, W. F. (1995). Small cysteine-rich antifungal proteins from radish: their role in host defense. *Plant Cell* **7**, 573 - 588.

Thoma, S., Kaneko, Y., and Somerville, C. (1993). A non-specific lipid transfer protein from *Arabidopsis* is a cell wall protein. *Plant J.* **3**, 427 - 436.

Thomma, B. P. H. J., Eggermont, K., Penninckx, I. A. M. A., Mauch-Mani, B., Vogelsang, R., Cammue, B. P. A., and Broekaert, W. F. (1998). Separate jasmonate-dependent and salicylate-dependent defense-response pathways in *Arabidopsis* are essential for resistance to distinct microbial pathogens. *Proc. Natl. Acad. Sci. USA* **95**, 15107-15111.

Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994). ClustalW: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalities and weight matrix choice. *Nucl. Acid. Res.* **22**, 4673 - 4680.

Tomayo, *et al.* (1999). Interpreting patterns of gene expression with self-organizing maps: methods and application to hematopoietic differentiation. *Proc. Natl. Acad. Sci. USA* 96, 2907-2912.

Towbin, H., Staehelin, T., and Gordon, J. (1979). Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc. Natl. Acad. Sci. USA* **76**, 4350 - 4354.

Tsuji, J., Zook, M., Hammerschmidt, R., Somerville, S., and Last, R. (1993). Evidence that tryptophan is not a direct biosynthetic intermediate of camalexin in *Arabidopsis thaliana*. *Physiol. Mol. Plant Pathol.* **43**, 221 - 229.

Uknes, S., Mauch-Mani, B., Moyer, M., Williams, S., Dincher, S., Chandler, D., Potter, S., Slusarenko, A., Ward, E., and Ryals, J. (1992). Acquired resistance in *Arabidopsis*. *Plant Cell* **4**, 645-656.

Uknes, S., Winter, A., Delaney, T., Vernooij, B., Morse, A., Friedrich, L., Potter, S., Ward, E., and Ryals, J. (1993). Biological induction of systemic acquired resistance in *Arabidopsis*. *Mol. Plant Microbe Interact.* **6**, 680 - 685.

Valvekens, D., Van Montagu, M., and Van Lijsebettens, M. (1988). *Agrobacterium tumefaciens*-mediated transformation of *Arabidopsis thaliana* root explants by using kanamycin selection. *Proc. Natl. Acad. Sci. USA* **85**, 5536-5540.

Van Loon, L. C. (1985). Pathogenesis-related proteins. *Plant Mol. Biol.* **4**, 111 - 116.

Van Loon, L. C., and Antoniw, J. F. (1982). Comparison of the effects of salicylic acid and ethephon with virus-induced hypersensitivity and acquired resistance in tobacco. *Neth. J. Pl. Path.* **88**, 237-256.

Van Wees, S. C. M., Pieterse, C. M. J., Trijssenaar, A., Van't Westende, Y. A. M., Hartog, F., and Van Loon, L. C. (1997). Differential induction of systemic resistance in *Arabidopsis* by biocontrol bacteria. *Mol. Plant Microbe Interact.* **10**, 716 - 724.

Vaux, D. L., and Strasser, A. (1996). The molecular biology of apoptosis. *Proc. Natl. Acad. Sci. USA* **93**, 2239 - 2244.

Velculescu, V. E., Zhang, L., Zhou, W., Vogelstein, J., Basrai, M. A., Bassett Jr., D. E., Hieter, P., Vogelstein, B., and Kinzler, K. W. (1997). Characterization of the yeast transcriptome. *Cell* **88**, 243 - 251.

Vernooij, B., Friedrich, L., Ahl Goy, P., Staub, T., Kessmann, H., and Ryals, J. (1995). 2,6-Dichloroisonicotinic acid-induced resistance to pathogens without the accumulation of salicylic acid. *Mol. Plant-Microbe Interact.* **8**, 228 - 234.

Vernooij, B., Friedrich, L., Morse, A., Reist, R., Kolditz-Jowhar, R., Ward, E., Uknes, S., Kessmann, H., and Ryals, J. (1994). Salicylic acid is not the translocated signal responsible for inducing systemic acquired resistance but is required in signal transduction. *Plant Cell* **6**, 959-965.

Vidal, S., Ponce de Leon, I., Denecke, J., and Palva, E. T. (1997). Salicylic acid and the plant pathogen *Erwinia carotovora* induce defense genes via antagonistic pathways. *Plant J.* **11**, 115 - 123.

Vignutelli, A., Wasternack, C., Apel, K., and Bohlmann, H. (1998). Systemic and local induction of an *Arabidopsis* thionin gene by wounding and pathogens. *Plant J.* **14**, 285-295.

Vijayan, P., Shockey, J., Levesque, C., Cook, R., and Browse, J. (1998). A role for jasmonate in pathogen defense of *Arabidopsis*. *Proc. Natl. Acad. Sci., USA* **95**, 7209-7214.

von Roepenack, E., Parr, A., and Schulze-Lefert, P. (1998). Structural analyses and dynamics of soluble and cell wall-bound phenolics in a broad-spectrum resistance to the powdery mildew fungus in barley. *J. Biol. Chem.* **273**.

Wanner, L. A., Li, G., Ware, D., Somssich, I. E., and Davis, K. R. (1995). The phenylalanine ammonia-lyase gene family in *Arabidopsis thaliana*. *Plant Mol. Biol.* **27**, 327 - 338.

Ward, E. R., Uknes, S. J., Williams, S. C., Dincher, S. S., Wiederhold, D. L., Alexander, D. C., Ahl-Goy, P., Métraux, J.-P., and Ryals, J. A. (1991). Coordinate gene activity in response to agents that induce systemic acquired resistance. *Plant Cell* **3**, 1085-1094.

Warren, R. F., Henk, A., Mowery, P., Holub, E., and Innes, R. W. (1998). A mutation within the leucin-rich repeat domain of the *Arabidopsis* disease resistance gene *RPS5* partially suppresses multiple bacterial and downy mildew resistance genes. *Plant Cell* **10**, 1439 - 1452.

Weymann, K., Hunt, M., Uknes, S., Neuenschwander, U., Lawton, K., Steiner, H-Y., and Ryals, J. (1995). Suppression and restoration of lesion formation in *Arabidopsis lsd* mutants. *Plant Cell* **7**, 2013 - 2022.

Whalen, M. C., Innes, R. W., Bent, A. F., and Staskawicz, B. J. (1991). Identification of *Pseudomonas syringae* pathogens of *Arabidopsis* and a bacterial locus determining avirulence on both *Arabidopsis* and soybean. *Plant Cell* **3**, 49-59.

White, R. F. (1979). Acetylsalicylic acid (aspirin) induces resistance to tobacco mosaic virus in tobacco. *Virology* **99**, 410-412.

Whitham, S., Dinesh-Kumar, S. P., Choi, D., Hehl, R., Corr, C., and Baker, B. (1994). The product of the tobacco mosaic virus resistance gene *N*: similarity to Toll and the interleukin-1 receptor. *Cell* **78**, 1101-1115.

Willits, M. G., and Ryals, J. A. (1998). Determining the relationship between salicylic acid levels and systemic acquired resistance induction in tobacco. *Mol. Plant Microbe Interact.* **11**, 795-800.

Winzeler, *et al.* (1999). Functional characterization of the S. cerevisiae genome by gene deletion and parallel analysis. *Science* 285: 901-906.

Wodicka, L., Dong, H., Mittmann, M., Ho, M. H., and Lockhart, D. J. (1997). Genome-wide expression monitoring in *Saccharomyces cerevisiae. Nat. Biotechnology* **15**, 1359 - 1367.

Xiao, S., Ellwood, S., Findlay, K., Olivier, R. P., and Turner, J. G. (1997). Characterization of three loci controlling resistance of *Arabidopsis thaliana* accession *Ms-0* to two powdery mildew diseases. *Plant J.* **12**, 757 - 768.

Yalpani, N., Leon, J., Lawton, M. A., and Raskin, I. (1993). Pathway of salicylic acid biosynthesis in healthy and virus-inoculated tobacco. *Plant Physiol.* **103**, 315-321.

Yang, Y., Shah, J., and Klessig, D. F. (1997). Signal perception and transduction in plant defense responses. *Genes Dev.* **11**, 1621 - 1639.

Yanish-Perron, C., Vieira, J., and Messing, J. (1985). Improved M13 phage cloning vectors and host strains: Nucleotide sequence of M13mp18 and pUC19 vectors. *Gene* **13**, 103 - 119.

Yao, K., De Luca, V., and Brisson, N. (1995). Creation of a metabolic sink for tryptophan alters the phenylpropanoid pathway and the susceptibility of potato to *Phytophtora infestans. Plant Cell* **7**, 1787 - 1799.

Yu, I.-C., Parker, J., and Bent, A. F. (1998). Gene-for-gene disease resistance without the hypersensitive response in *Arabidopsis dnd1* mutant. *Proc. Natl. Acad. Sci. USA* **95**, 7819-7824.

Zhao, J., and Last, R. L. (1996). Coordinate regulation of the tryptophan biosynthetic pathway and indolic phytoalexin accumulation in *Arabidopsis. Plant Cell* **8**, 2235 - 2244.

What is claimed is:

1. A method for isolating a regulatory DNA sequence from a differentially expressed gene, comprising:

(a) obtaining an expression profile of a genome under two or more different conditions;

(b) comparing gene induction and repression patterns among each of the different conditions;

(c) identifying a gene that is differentially expressed among the different conditions;

(d) obtaining the sequence of genomic DNA comprising the gene identified as being differentially expressed and regulatory elements associated with said gene; and

(e) isolating a regulatory DNA sequence associated with the differentially expressed gene from the genomic DNA.

2. A method for isolating a common regulatory DNA sequence from a group of co-regulated genes, comprising:

(a) obtaining an expression profile of a genome under two or more different conditions;

(b) comparing gene induction and repression patterns among each of the different conditions;

(c) identifying a group of genes that are co-regulated with respect to each other but that are differentially expressed with respect to the rest of the genome among the different conditions;

(d) obtaining the sequence of genomic DNA comprising the genes identified as being differentially expressed and regulatory elements associated with said genes;

(e) identifying common regulatory DNA sequences among the genomic DNA associated with the co-regulated genes; and

(f) isolating a common regulatory DNA sequence associated with the differentially expressed genes from the genomic DNA.

3. A method according to claim 1 or claim 2, wherein said regulatory DNA sequence is a promoter.

4. A method according to claim 1 or claim 2, wherein said genome is a plant genome.

5. A method according to claim 1 or claim 2, wherein said expression profile is obtained using a DNA microarray.

6. A method according to claim 1 or claim 2, wherein said two or more different conditions comprise biotic stimuli.

7. A method according to claim 1 or claim 2, wherein said two or more different conditions comprise abiotic stimuli.

8. A method according to claim 1 or claim 2, wherein said two or more different conditions comprise biotic and abiotic stimuli.

9. A method according to claim 1 or claim 2, wherein said genome is a plant genome and wherein said two or more different conditions comprise at least one SAR inducing condition or at least one SAR repressing condition.

10. A method according to claim 9, wherein said least one SAR inducing condition comprises pathogen infection, SA application, BTH application, *NIM1* gene expression, or a *cim* mutation.

11. A method according to claim 9, wherein said least one SAR repressing condition comprises NahG expression or a *nim1* mutation.

12. A method for isolating a regulatory DNA sequence from a differentially expressed plant gene, comprising:

(a) using DNA microarray technology to obtain an expression profile of a plant genome under two or more different conditions, wherein said conditions comprise at least one SAR-inducing or SAR-repressing condition;

(b) comparing gene induction and repression patterns among each of the different conditions;

(c) identifying a gene that is differentially expressed among the different conditions;
(d) obtaining the sequence of genomic DNA comprising the gene identified as being differentially expressed and regulatory elements associated with said gene; and
(e) isolating a regulatory DNA sequence associated with the differentially expressed gene from the genomic DNA.

13. A method for isolating a common regulatory DNA sequence from a group of co-regulated plant genes, comprising:
(a) using DNA microarray technology to obtain an expression profile of a genome under two or more different conditions, wherein said conditions comprise at least one SAR-inducing or SAR-repressing condition;
(b) comparing gene induction and repression patterns among each of the different conditions;
(c) identifying a group of genes that are co-regulated with respect to each other but that are differentially expressed with respect to the rest of the genome among the different conditions;
(d) obtaining the sequence of genomic DNA comprising the genes identified as being differentially expressed and regulatory elements associated with said genes;
(e) identifying common regulatory DNA sequences among the genomic DNA associated with the co-regulated genes; and
(f) isolating a common regulatory DNA sequence associated with the differentially expressed genes from the genomic DNA.

14. A method according to claim 12 or claim 13, wherein said regulatory DNA sequence is a promoter.

15. A method according to claim 12 or claim 13, wherein said least one SAR inducing condition comprises pathogen infection, SA application, BTH application, *NIM1* gene expression, or a *cim* mutation.

16. A method according to claim 12 or claim 13, wherein said least one SAR repressing condition comprises NahG expression or a *nim1* mutation.

17. A method according to claim 12 or claim 13, wherein said two or more different conditions comprise BTH application, and at least one condition selected from the group consisting of pathogen infection, SA application, and a *nim1* mutation.

18. A method according to claim 17, wherein said two or more different conditions comprise BTH application, pathogen infection, and SA application, and wherein one or more genes are identified that are inducible by BTH application but not by pathogen infection or SA application.

19. A method according to claim 17, wherein said two or more different conditions comprise BTH application and a *nim1* mutation, and wherein one or more genes are identified that are inducible by BTH application in *nim1* mutant plants.

20. A method according to claim 18 or claim 19, wherein said one or more genes are inducible at least 5x by BTH application.

SEQUENCE LISTING

<110> Maleck, Klaus
Lawton, Kay
Dietrich, Robert

<120> IDENTIFICATION OF DNA CONTROL ELEMENTS RESPONSIVE TO SPECIFIC STIMULI

<130> A-30901A

<140>
<141>

<160> 8

<170> PatentIn Ver. 2.1

<210> 1
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
oligonucleotide primer

<400> 1
taatacgact cactataggg 20

<210> 2
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
oligonucleotide primer

<400> 2
atttaggtga cactatagga 20

<210> 3
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
oligonucleotide primer

<400> 3
ttcccaacag ttgcgcagc 19

<210> 4
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
oligonucleotide primer

<400> 4
tcttcgctat tacgccagct 20

<210> 5
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
oligonucleotide primer

<400> 5
tcacacagga aacagctat 19

<210> 6
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
oligonucleotide primer

<400> 6
acacaacata cgagccggaa 20

<210> 7
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
oligonucleotide primer

<400> 7
tgtaaaacga cggccagt 18

<210> 8
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
oligonucleotide primer

<400> 8
ggaaacagct atgaccat 18

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
2 November 2000 (02.11.2000)

PCT

(10) International Publication Number
WO 00/65039 A3

(51) International Patent Classification[7]: C12N 15/82

(21) International Application Number: PCT/US00/11460

(22) International Filing Date: 28 April 2000 (28.04.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/131,433 28 April 1999 (28.04.1999) US
60/144,669 19 July 1999 (19.07.1999) US

(71) Applicant *(for all designated States except US)*: **NOVARTIS AG** [CH/CH]; Schwarzwaldallee 215, D-4058 Basel (CH).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **MALECK, Klaus** [DE/DE]; Augustenstrasse 74, D-80333 München (DE). **LAWTON, Kay, Ann** [US/US]; 8016 Crooked Chute Court, Raleigh, NC 27612 (US). **DIETRICH, Robert, Arthur** [US/US]; 25 Copper Hill Court, Durham, NC 27713 (US).

(74) Agent: **MEIGS, J., Timothy**; Novartis corporation, Patent & Trademark Dept., 3054 Cornwallis Road, Research Triangle Park, NC 27709-2257 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— *With international search report.*
— *Before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments.*

(88) Date of publication of the international search report:
8 February 2001

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: IDENTIFICATION OF DNA CONTROL ELEMENTS RESPONSIVE TO SPECIFIC STIMULI

(57) Abstract: In order to identify genes whose proteins are involved in the regulation of SAR downstream of cell death, a screen was carried out to isolate mutants that constitutively express SAR. The *PR-1* promoter, the most reliable marker for the onset of SAR, was cloned in front of the firefly *(Photinus pyralis)* luciferase reporter gene and transformed into *Arabidopsis*. A transgenic line with a luciferase expression pattern that paralleled expression of the endogenous *PR-1* gene was identified and subjected to EMS mutagenesis to isolate mutants that constitutively expressed the *PR-1*/luciferase gene. Biochemical, cytological, pathological, and genetic methods were employed to further characterize the mutants and to prove the isolation of SAR mutants downstream of cell death. This characterization enables one to distinguish between classes of *cim* mutants, and to describe physiological changes that occur during the maintenance phase of SAR. Furthermore, utilizing microarray chips, the entire plant genome was simultaneously surveyed for genes that change in expression in response to biotic and abiotic factors. By comparing gene expression changes across various treatments, groups of co-regulated genes (regulons) were identified and the genomic sequences of genes within a regulon were examined to identify common sequence motifs that are likely to act as regulatory elements. This approach used experimental design based upon the biology of the study system in combination with bioinformatics to analyze the results.

WO 00/65039 A3

INTERNATIONAL SEARCH REPORT

Intern. nal Application No
PCT/US 00/11460

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12N15/82

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, WPI Data, PAJ, BIOSIS, MEDLINE, CHEM ABS Data, EMBASE, SCISEARCH, BIOTECHNOLOGY ABS

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	SPELLMAN ET AL: "COMPREHENSIVE IDENTIFICATION OF CELL CYCLE-REGULATED GENES OF THE YEAST SACCHAROMYCES CEREVISIAE BY MICROARRAY HYBRIDIZATION" MOLECULAR BIOLOGY OF THE CELL,US,BETHESDA, MD, vol. 9, December 1998 (1998-12), pages 3273-3297, XP002143737 ISSN: 1059-1524 figures 1-6 page 3279, right-hand column, paragraph 2; figure 2; table 4 --- -/--	1-3,5-8

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

'A' document defining the general state of the art which is not considered to be of particular relevance

'E' earlier document but published on or after the international filing date

'L' document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

'O' document referring to an oral disclosure, use, exhibition or other means

'P' document published prior to the international filing date but later than the priority date claimed

'T' later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

'X' document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

'Y' document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

'&' document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
17 November 2000	01/12/2000

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	ALCONADA RODRIG.., A

3

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Intern. nal Application No
PCT/US 00/11460

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	SOMERVILLE, S. C. ET AL: "Alternate methods of gene discovery - the candidate EST approach and DNA microarrays" CELLULAR INTEGRATION OF SIGNALLING PATHWAYS IN PLANT DEVELOPMENT. PROCEEDINGS OF THE NATO ADVANCED STUDY INSTITUTE, ACQUA FREDDA DI MARATEA, ITALY, 20-30 MAY, 1997, (1998) PP. 297-309. NATO ASI SERIES. SERIES H: CELL BIOLOGY VOL. 104., XP000964578 Department of Plant Biology, Carnegie Institution of Washington, Stanford, CA 94305, USA page 305 -page 306, paragraph 1; figure 2; table 1 ---	1-20
Y	US 5 851 766 A (HARMS CHRISTIAN ET AL) 22 December 1998 (1998-12-22) column 22, line 12-55 column 33, line 27-52 column 34, line 40-55 examples 8,28,33,34 SEQ ID NO:1 ---	1-6,9-16
Y	MOLINA ANTONIO ET AL: "Impaired fungicide activity in plants blocked in disease resistance signal transduction." PLANT CELL, vol. 10, no. 11, November 1998 (1998-11), pages 1903-1914, XP000885732 ISSN: 1040-4651 page 1903, right-hand column, paragraph 2 page 1904, left-hand column, paragraph 2 page 1904, left-hand column, paragraph 4 ---	7-11, 15-20
A	SCHENA M ET AL: "QUANTITATIVE MONITORING OF GENE EXPRESSION PATTERNS WITH A COMPLEMENTARY DNA MICROARRAY" SCIENCE,US,AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE,, vol. 270, no. 5235, 20 October 1995 (1995-10-20), pages 467-470, XP000644675 ISSN: 0036-8075 cited in the application the whole document --- -/--	1-20

3

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Intern. nal Application No
PCT/US 00/11460

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document. with indication,where appropriate, of the relevant passages	Relevant to claim No.
A	DESPREZ THIERRY ET AL: "Differential gene expression in Arabidopsis monitored using cDNA arrays." PLANT JOURNAL, vol. 14, no. 5, June 1998 (1998-06), pages 643-652, XP000960485 ISSN: 0960-7412 cited in the application figure 5; table 1 page 650, right-hand column, paragraph 3 ---	1-20
A	RUAN YIJUN ET AL: "Towards Arabidopsis genome analysis: Monitoring expression profiles of 1400 genes using cDNA microarrays." PLANT JOURNAL, vol. 15, no. 6, 1998, pages 821-833, XP000960486 ISSN: 0960-7412 cited in the application figures 4-6; tables 1,2 ---	1-20
A	RAFALSKI J ANTONI ET AL: "New experimental and computational approaches to the analysis of gene expression." ACTA BIOCHIMICA POLONICA, vol. 45, no. 4, 1998, pages 929-934, XP000952158 ISSN: 0001-527X page 933, right-hand column, paragraph 5 ---	1-20
A	WARD E R ET AL: "COORDINATE GENE ACTIVITY IN RESPONSE TO AGENTS THAT INDUCE SYSTEMIC ACQUIRED RESISTANCE" PLANT CELL, vol. 3, no. 10, 1991, pages 1085-1094, XP002134008 ISSN: 1040-4651 cited in the application page 1087, left-hand column -page 1089, right-hand column, paragraph 1; figures 2,3 ---	1-20
T	DE 198 60 313 A (FELSENSTEIN FRIEDRICH ;THUEMMLER FRITZ (DE)) 29 June 2000 (2000-06-29) page 4, line 20,21; claim 9; figure 1; example 1 --- -/--	1-20

3

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Intern. nal Application No
PCT/US 00/11460

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
T	RAMONELL K M ET AL: "Microarray dissection of systemic acquired resistance in Arabidopsis plants infected with powdery mildew." PHYTOPATHOLOGY, vol. 90, no. 6 Supplement, June 2000 (2000-06), page S63 XP000964562 Annual Meeting of the American Phytopathological Society;New Orleans, Louisiana, USA; August 12-16, 2000 ISSN: 0031-949X abstract ---	1-20
T	SCHENK P M ET AL: "Coordinated plant defense responses in Arabidopsis revealed by microarray analysis" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA, (10 OCT 2000) VOL. 97, NO. 21, PP. 11655-11660. PUBLISHER: NATL ACAD SCIENCES, 2101 CONSTITUTION AVE NW, WASHINGTON, DC 20418. ISSN: 0027-8424., XP002153163 the whole document -----	1-20

3

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Information on patent family members

Internat. nal Application No
PCT/US 00/11460

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 5851766 A	22-12-1998	NONE	
DE 19860313 A	29-06-2000	NONE	